# Notes for *Foundations of Modern Analysis* by Avner Friedman

Anton Ottosson – antonott@kth.se

May 12, 2025

# Chapter 1 − Measure Theory

## Section 1.1 − Rings and Algebras

### Problems

**1.1.1**

$$\left(\varliminf_n E_n\right)^c = \varlimsup_n E_n^c, \quad \left(\varlimsup_n E_n\right)^c = \varliminf_n E_n^c.$$

*Solution.* Note that

$$x \in \varliminf_n E_n \iff x \in E_n \text{ for all but finitely many } n$$

$$\iff x \in E_n^c \text{ for only finitely many } n.$$

Hence

$$x \in \left(\varliminf_n E_n\right)^c \iff x \in E_n^c \text{ for infinitely many } n$$

$$\iff x \in \varlimsup_n E_n^c,$$

proving the first identity.

Next, let $F_n = E_n^c$ for every $n$. Then

$$\varlimsup_n E_n = \varlimsup_n F_n^c = \left(\varliminf_n F_n\right)^c = \left(\varliminf_n E_n^c\right)^c$$

by the first identity, and the second identity follows.

**1.1.2**

$$\varlimsup_n E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n, \quad \varliminf_n E_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n.$$

*Solution.* Suppose $x \in \overline{\lim}_n E_n$. Then $x \in E_n$ for infinitely many $n$. It follows that $x \in \bigcup_{n=k}^{\infty} E_n$ for all $k \in \mathbb{N}$, and hence that $x \in \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n$.

Conversely, assume that $x \in \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n$. Then $x \in \bigcup_{n=k}^{\infty} E_n$ for all $k \in \mathbb{N}$. It follows that $x \in E_n$ for infinitely many $n$, and thus that $x \in \overline{\lim}_n E_n$. This proves the first identity.

Next, suppose that $x \in \underline{\lim}_n E_n$. Then $x \in E_n$ for all but finitely many $n$, so there is some $k' \in \mathbb{N}$ such that $x \in E_n$ for all $n \geq k'$. It follows that $x \in \bigcap_{n=k'}^{\infty} E_n$, and hence that $x \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n$.

Conversely, assume that $x \in \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n$. Then $x \in \bigcap_{n=k'}^{\infty} E_n$ for some $k' \in \mathbb{N}$, which means that $x \in E_n$ for all $n \geq k'$. It follows that $x \in E_n$ for all but finitely many $n$; that is, $x \in \underline{\lim}_n E_n$.

### 1.1.3

If $\mathscr{R}$ is a $\sigma$-ring and $E_n \in \mathscr{R}$, then

$$\bigcap_{n=1}^{\infty} E_n \in \mathscr{R}, \quad \overline{\lim_n} E_n \in \mathscr{R}, \quad \underline{\lim_n} E_n \in \mathscr{R}.$$

*Solution.* Note that

$$\bigcap_{n=1}^{\infty} E_n = E_1 \cap \left( \bigcap_{n=1}^{\infty} E_n \right) = E_1 - \left( E_1 - \bigcap_{n=1}^{\infty} E_n \right),$$

and that

$$E_1 - \bigcap_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} (E_1 - E_n) \in \mathscr{R},$$

using one of De Morgan's laws along with properties (b) and (e). It follows (by (b) again) that $\bigcap_{n=1}^{\infty} E_n \in \mathscr{R}$. That is, $\mathscr{R}$ is closed under countable intersections.

Given $k \in \mathbb{N}$, let $A_n = \varnothing$ for $n < k$, and let $A_n = E_n$ for $n \geq k$. Then $A_n \in \mathscr{R}$ for all $n$ by (a), hence

$$\bigcup_{n=k}^{\infty} E_n = \bigcup_{n=1}^{\infty} A_n \in \mathscr{R}$$

by (e). Since $\mathscr{R}$ is closed under countable intersections, we then have

$$\overline{\lim_n} E_n = \bigcap_{k=1}^{\infty} \bigcup_{n=k}^{\infty} E_n \in \mathscr{R}.$$

By a similar argument we find that

$$\bigcap_{n=k}^{\infty} E_n \in \mathscr{R}$$

2

for all $k \in \mathbb{N}$. Thus

$$\varliminf_n E_n = \bigcup_{k=1}^{\infty} \bigcap_{n=k}^{\infty} E_n \in \mathscr{R}$$

by (e).

### 1.1.4

The intersection of any collection of rings (algebras, $\sigma$-rings, or $\sigma$-algebras) is also a ring (an algebra, $\sigma$-ring, or $\sigma$-algebra).

*Solution.* Let $\mathscr{C}$ be a collection of classes. Let $\bigcap \mathscr{C}$ denote the intersection of all classes in $\mathscr{C}$. We will show that if one of the properties (a)-(e) is satisfied by all classes in $\mathscr{C}$, then $\bigcap \mathscr{C}$ satisfies that property as well. The result requested in the problem then follows as an immediate corollary.

It is clear that if every $\mathscr{R} \in \mathscr{C}$ satisfies (a), then so does $\bigcap \mathscr{C}$. Suppose every $\mathscr{R} \in \mathscr{C}$ satisfies (b). If $A, B \in \bigcap \mathscr{C}$ then $A, B \in \mathscr{R}$ for every $\mathscr{R} \in \mathscr{C}$. Hence $A - B \in \mathscr{R}$ for all $\mathscr{R} \in \mathscr{C}$, and it follows that $A - B \in \bigcap \mathscr{C}$. The argument for (c) is similar (with $A \cup B$ in place of $A - B$), and (d) is obvious.

Finally, suppose that every $\mathscr{R} \in \mathscr{C}$ satisfies (e). If $A_1, A_2, \ldots \in \bigcap \mathscr{C}$ then $A_1, A_2, \ldots \in \mathscr{R}$ for every $\mathscr{R} \in \mathscr{C}$. Hence $\bigcup_{n=1}^{\infty} A_n \in \mathscr{R}$ for all $\mathscr{R} \in \mathscr{C}$, and it follows that $\bigcup_{n=1}^{\infty} A_n \in \bigcap \mathscr{C}$.

### 1.1.5

If $\mathscr{D}$ is any class of sets, then there exists a unique ring $\mathscr{R}_0$ such that (i) $\mathscr{R}_0 \supset \mathscr{D}$, and (ii) any ring $\mathscr{R}$ containing $\mathscr{D}$ contains also $\mathscr{R}_0$. $\mathscr{R}_0$ is called the *ring generated* by $\mathscr{D}$, and is denoted by $\mathscr{R}(\mathscr{D})$.

*Solution.* Let $\mathscr{R}_0$ be the intersection of all rings containing $\mathscr{D}$. This is a ring by the previous exercise, and it satisfies the properties (i) and (ii). To see that it is unique, let $\mathscr{R}_0'$ also by a ring satisfying (i) and (ii). Then $\mathscr{R}_0 \subset \mathscr{R}_0'$ and $\mathscr{R}_0' \subset \mathscr{R}_0$ by property (ii), so $\mathscr{R}_0 = \mathscr{R}_0'$.

### 1.1.6

If $\mathscr{D}$ is any class of sets, then there exists a unique $\sigma$-ring $\mathscr{S}_0$ such that (i) $\mathscr{S}_0 \supset \mathscr{D}$, and (ii) any $\sigma$-ring containing $\mathscr{D}$ contains also $\mathscr{S}_0$. We call $\mathscr{S}_0$ the $\sigma$-*ring generated* by $\mathscr{D}$, and denote it by $\mathscr{S}(\mathscr{D})$. A similar result holds for $\sigma$-algebras, and we speak of the $\sigma$-*algebra generated* by $\mathscr{D}$.

*Solution.* By the same argument as in the previous exercise, $\mathscr{S}_0$ is the intersection of all $\sigma$-rings containing $\mathscr{D}$. Similarly the $\sigma$-algebra generated by $\mathscr{D}$ is the intersection of all $\sigma$-algebras containing $\mathscr{D}$.

### 1.1.7

If $\mathscr{D}$ is any class of sets, then every set in $\mathscr{R}(\mathscr{D})$ can be covered by (that is, is contained in) a finite union of sets of $\mathscr{D}$. [*Hint:* The class $\mathscr{K}$ of sets that can be covered by finite unions of sets of $\mathscr{D}$ forms a ring.]

*Solution.* Let $\mathscr{K}$ be the class of all sets that can be covered by a finite union of sets in $\mathscr{D}$. Certainly $\varnothing \in \mathscr{K}$, since $\varnothing$ is a subset of the empty union. If $A, B \in \mathscr{K}$, then

$$A \subset \bigcup_{i=1}^{m} E_i, \quad B \subset \bigcup_{i=1}^{n} F_i,$$

for some sets $E_1, \ldots, E_m, F_1, \ldots, F_n \in \mathscr{D}$. (Note that $m$ or $n$ can be zero, in which case the corresponding union is empty.) Thus

$$A - B \subset A \subset \bigcup_{i=1}^{m} E_i$$

and

$$A \cup B \subset \left( \bigcup_{i=1}^{m} E_i \right) \cup \left( \bigcup_{j=1}^{n} F_j \right),$$

so both $A - B$ and $A \cup B$ are elements of $\mathscr{K}$.

The above shows that $\mathscr{K}$ is a ring, and certainly $\mathscr{D} \subset \mathscr{K}$. Hence $\mathscr{R}(\mathscr{D}) \subset \mathscr{K}$ by Problem 1.1.5, and it follows that every set in $\mathscr{R}(\mathscr{D})$ can be covered by a finite union of sets in $\mathscr{D}$.

## Section 1.2 – Definition of Measure

### Problems

#### 1.2.1

If $\mu$ satisfies the properties (i)-(iii) in Definition 1.2.1, and if $\mu(E) < \infty$ for at least one set $E$, then (iv) is also satisfied.

*Solution.* We have

$$\mu(E) = \mu(E \cup \varnothing) = \mu(E) + \mu(\varnothing),$$

hence $\mu(\varnothing) = 0$.

#### 1.2.2

Let $X$ be an infinite space. Let $\mathcal{A}$ be the class of all subsets of $X$. Define $\mu(E) = 0$ if $E$ is finite and $\mu(E) = \infty$ if $E$ is infinite. Then $\mu$ is finitely additive but not completely additive.

*Solution.* Suppose $A, B \in \mathcal{A}$. Note that $A \cup B$ is finite if both $A$ and $B$ are finite, but infinite otherwise. Hence

$$\mu(A \cup B) = 0 = \mu(A) + \mu(B)$$

in the former case, and

$$\mu(A \cup B) = \infty = \mu(A) + \mu(B)$$

in the latter. This proves that $\mu$ is additive; *finite* additivity follows by a simple induction argument.

Let $(x_n)$ be a sequence of distinct points in $X$. Then $\bigcup_{n=1}^{\infty}\{x_n\}$ is an infinite set, so

$$\mu\left(\bigcup_{n=1}^{\infty}\{x_n\}\right) = \infty,$$

but

$$\sum_{n=1}^{\infty}\mu(\{x_n\}) = 0.$$

Thus $\mu$ is not completely additive.

### 1.2.3

If $\mu$ is a measure on a $\sigma$-algebra $\mathcal{A}$, and if $E$, $F$ are sets of $\mathcal{A}$, then

$$\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F).$$

*Solution.* If $\mu(F) = \infty$, then $\mu(E \cup F) = \infty$ by Theorem 1.2.1(i), and the given equality holds. If $\mu(F) < \infty$, then

$$\begin{aligned}
\mu(E \cup F) &= \mu[E \cup (F - (E \cap F))] \\
&= \mu(E) + \mu[F - (E \cap F)] \\
&= \mu(E) + \mu(F) - \mu(E \cap F),
\end{aligned}$$

with the last equality following from Theorem 1.2.1(ii). Note that $E \cap F \subset F$ so that $\mu(E \cap F) \leq \mu(F) < \infty$. Hence we can rearrange the above to yield

$$\mu(E) + \mu(F) = \mu(E \cup F) + \mu(E \cap F).$$

### 1.2.6

Give an example of a measure $\mu$ and a monotone-decreasing sequence $\{E_n\}$ of $\mathcal{A}$ such that $\mu(E_n) = \infty$ for all $n$, and $\mu(\lim_n E_n) = 0$.

*Solution.* Let $X = \mathbb{R}$ and let $\mathcal{A} = \mathcal{P}(\mathbb{R})$ (the power set of $\mathbb{R}$; this is easily seen to be a $\sigma$-algebra). Define $\mu : \mathcal{A} \to [0, \infty]$ such that $\mu(E)$ is the number of points in $E$ (with $\mu(E) = \infty$ if $E$ is infinite). This is easily seen to be a measure.

For each $n \in \mathbb{N}$, let $E_n = (0, 1/n)$. Then $(E_n)$ is a monotone decreasing sequence of sets in $\mathcal{A}$, $\mu(E_n) = \infty$ for all $n$, and

$$\mu\left(\lim_n E_n\right) = \mu\left(\bigcap_{n=1}^{\infty} E_n\right) = \mu(\varnothing) = 0.$$

## Section 1.3 – Outer Measure

### Problems

#### 1.3.1

Define $\mu^*(E)$ as the number of points in $E$ if $E$ is finite and $\mu^*(E) = \infty$ if $E$ is infinite. Show that $\mu^*$ is an outer measure. Determine the measurable sets.

*Solution.* Of the properties listed in Definition 1.3.1, only countable subadditivity is non-obvious for $\mu^*$. But let us start with proving finite subadditivity.

Let $A$ and $B$ be sets. If either is infinite, then so is $A \cup B$, hence

$$\mu^*(A \cup B) = \infty = \mu^*(A) + \mu^*(B).$$

If both $A$ and $B$ are finite sets, then

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B - A) \le \mu^*(A) + \mu^*(B)$$

by basic set-theoretic considerations. Thus $\mu^*$ is subadditive, and finite subadditivity follows by induction on the number of sets in the union.

Now, let $(E_n)$ be a sequence of sets. If infinitely many of the sets $E_n$ are nonempty, then $\sum_n \mu^*(E_n) = \infty$, and

$$\mu^* \left( \bigcup_n E_n \right) \le \sum_n \mu^*(E_n)$$

follows. If only finitely many of the sets $E_n$ are nonempty, let $E_{n_1}, E_{n_2}, \ldots, E_{n_k}$ be those sets. Then

$$\mu^* \left( \bigcup_{n=1}^{\infty} E_n \right) = \mu^* \left( \bigcup_{i=1}^{k} E_{n_i} \right) \le \sum_{i=1}^{k} \mu^*(E_{n_i}) = \sum_{n=1}^{\infty} \mu^*(E_n),$$

by finite subadditivity. This proves that $\mu^*$ is countably subadditive, and hence that $\mu^*$ is an outer measure.

Note that $\mu^*$ is *additive* on disjoint sets; if $A \cap B = \varnothing$, then $\mu^*(A \cup B) = \mu^*(A) + \mu^*(B)$. In particular,

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A - E)$$

for all sets $A, E$. That is, all sets are measurable.

#### 1.3.2

Define $\mu^*(\varnothing) = 0$, $\mu^*(E) = 1$ if $E \ne \varnothing$. Show that $\mu^*$ is an outer measure, and determine the measurable sets.

*Solution.* As in the previous excercise, the only slightly non-obvious property is countable subadditivity. Hence, let $(E_n)$ be a sequence of sets. If all the sets $E_n$ are empty, then certainly

$$\mu^* \left( \bigcup_n E_n \right) = 0 = \sum_n \mu^*(E_n).$$

If not, then there is some $m$ such that $E_m \neq \varnothing$, and it follows that

$$\mu^* \left( \bigcup_n E_n \right) = 1 = \mu^*(E_m) \leq \sum_n \mu^*(E_n).$$

Thus $\mu^*$ is indeed countably subadditive, and therefore also an outer measure.

The empty set is measurable:

$$\mu^*(A \cap \varnothing) + \mu^*(A - \varnothing) = \mu^*(\varnothing) + \mu^*(A) = \mu^*(A)$$

for all sets $A$. It follows by Theorem 1.3.1 that $X$ is measurable as well (the measurable sets make up a $\sigma$-algebra). Indeed $\varnothing$ and $X$ are the only measurable sets. To see this, let $E$ be any set other than those two (this requires that $X$ contains at least two elements). Then both $E$ and $E^c$ are nonempty, so

$$\mu^*(X \cap E) + \mu^*(X - E) = \mu^*(E) + \mu^*(E^c) = 2 > 1 = \mu^*(X).$$

# Section 1.4 − Construction of Outer Measures

## Problems

### 1.4.4

If $\mathscr{K}$ is a $\sigma$-algebra and $\lambda$ is a measure on $\mathscr{K}$, then $\mu^*(A) = \lambda(A)$ for any $A \in \mathscr{K}$. [*Hint:* $\mu^*(A) = \inf\{\lambda(E); E \in \mathscr{K}, E \supset A\}$.]

*Solution.* Note that the description of $\mu^*$ can be simplified when $\mathscr{K}$ is a $\sigma$-algebra and $\lambda$ is a measure. For suppose that $A \subset X$, $E_n \in \mathscr{K}$ $(n = 1, 2, \dots)$, and $A \subset \bigcup_n E_n$. Then $E := \bigcup_n E_n \in \mathscr{K}$, and $\lambda(E) \leq \sum_n \lambda(E_n)$ by Theorem 1.2.2. Hence

$$\mu^*(A) = \inf\{\lambda(E); E \in \mathscr{K}, E \supset A\}.$$

Now, suppose that $A \in \mathscr{K}$. Certainly $\lambda(A)$ is an element of $\{\lambda(E); E \in \mathscr{K}, E \supset A\}$. And if $E \in \mathscr{K}$ and $E \supset A$, then $\lambda(E) \geq \lambda(A)$ by Theorem 1.2.1(i). Thus

$$\lambda(A) = \inf\{\lambda(E); E \in \mathscr{K}, E \supset A\} = \mu^*(A).$$

**1.4.5**

If $\mathscr{K}$ is a $\sigma$-algebra and $\lambda$ is a measure on $\mathscr{K}$, then every set in $\mathscr{K}$ is $\mu^*$-measurable.

*Solution.* Recall the simplified description of $\mu^*$ from the previous problem. Let $E \in \mathscr{K}$ and $A \subset X$. For every $\epsilon > 0$ there exists $F \in \mathscr{K}$ such that $F \supset A$ and

$$\mu^*(A) + \epsilon > \lambda(F);$$

else $\mu^*(A)$ would not be the greatest lower bound of $\{\lambda(E); E \in \mathscr{K}, E \supset A\}$. Moreover,

$$\lambda(F) = \lambda(F \cap E) + \lambda(F - E)$$

since $\lambda$ is a measure on $\mathscr{K}$,

$$\lambda(F \cap E) + \lambda(F - E) = \mu^*(F \cap E) + \mu^*(F - E)$$

by what we found in the previous exercise, and finally

$$\mu^*(F \cap E) + \mu^*(F - E) \geq \mu^*(A \cap E) + \mu^*(A - E)$$

by monotonicity of the outer measure $\mu^*$. Putting all of this together, we have

$$\mu^*(A) + \epsilon > \mu^*(A \cap E) + \mu^*(A - E)$$

for all $\epsilon > 0$, and thus

$$\mu^*(A) \geq \mu^*(A \cap E) + \mu^*(A - E).$$

It follows that every set $E \in \mathscr{K}$ is $\mu^*$-measurable.

# Section 1.6 – The Lebesgue and the Lebesgue-Stieltjes Measures

## Problems

**1.6.3**

The outer Lebesgue measure of a closed bounded interval $[a, b]$ on the real line is equal to $b - a$. [*Hint:* Use the Heine-Borel theorem to replace a countable covering by a finite covering.]

*Solution.* Suppose $(E_n)$ is a sequence of elements of $\mathscr{K}$ (i.e. a sequence of open intervals) such that $[a, b] \subset \bigcup_{n=1}^{\infty} E_n$. The collection $\{E_n\}$ constitutes an *open cover* of $[a, b]$. By the Heine-Borel theorem $[a, b]$ is compact, hence there exists a *finite subcover* $\{E_{n_1}, \ldots, E_{n_k}\}$, such that $[a, b] \subset \bigcup_{i=1}^{k} E_{n_i}$.

Assume without loss of generality that $E_{n_i} \cap [a, b] \neq \varnothing$ for all $i$; otherwise we can simply remove those $E_{n_i}$ that are disjoint with $[a, b]$ and still have a finite subcover. Write $E_{n_i} = (a_i, b_i)$ for each $i$, and define

$$\alpha = \min\{a_1, \ldots, a_k\}, \quad \beta = \max\{b_1, \ldots, b_k\}.$$

It is clear that $\alpha$ and $\beta$ are the infimum and supremum, respectively, of $\bigcup_{i=1}^{k} E_{n_i}$. Note that $\alpha = a_j$ for some $j$, and $a_j < a < b_j$ since $E_{n_j}$ and $[a, b]$ have nonempty intersection. Thus $(\alpha, a] \subset \bigcup_{i=1}^{k} E_{n_i}$, and similarly $[b, \beta) \subset \bigcup_{i=1}^{k} E_{n_i}$. It follows that

$$\bigcup_{i=1}^{k} E_{n_i} = (\alpha, \beta) \in \mathscr{K}.$$

Finally note that $\lambda$ is finitely subadditive. (This is easily proven with induction.) (TODO: This is not convincing; use better proof from Rosenthal notes.) Thus,

$$\sum_{n=1}^{\infty} \lambda(E_n) \geq \sum_{i=1}^{k} \lambda(E_{n_i}) \geq \lambda\left[(\alpha, \beta)\right] = \beta - \alpha > b - a.$$

It follows that $b - a$ is a lower bound of the set

$$\Lambda([a, b]) := \left\{ \sum_{n=1}^{\infty} \lambda(E_n); \ E_n \in \mathscr{K}, \ \bigcup_{n=1}^{\infty} E_n \supset [a, b] \right\}.$$

Moreover, for every $\epsilon > 0$ we have

$$[a, b] \subset \left( a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2} \right) \in \mathscr{K}$$

and

$$\lambda\left[ \left( a - \frac{\epsilon}{2}, b + \frac{\epsilon}{2} \right) \right] = b - a + \epsilon.$$

Hence $b - a$ is the *greatest* lower bound of $\Lambda([a, b])$, and $\mu^*([a, b]) = b - a$.

### 1.6.4

The outer Lebesgue measure of each of the intervals $(a, b), [a, b), (a, b]$ is equal to $b - a$.

*Solution.* Recall that $\mu^*$ is monotone, on account of being an outer measure. Hence $\mu^*[(a, b)] \leq \mu^*([a, b]) = b - a$, the latter equality being the result of the previous problem. Moreover, for all $\epsilon \in (0, b - a)$ we have

$$\left[ a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2} \right] \subset (a, b),$$

so that

$$\mu^*[(a, b)] \geq \mu^* \left( \left[ a + \frac{\epsilon}{2}, b - \frac{\epsilon}{2} \right] \right) = b - a - \epsilon.$$

Thus $\mu^*[(a, b)] \geq b - a$, and it follows that $\mu^*[(a, b)] = b - a$.

The outer measures of $[a, b)$ and $(a, b]$ follow immediately by monotonicity:

$$\mu^*[(a, b)] \leq \mu^*([a, b)) \leq \mu^*([a, b]),$$

so that $\mu^*([a, b)) = b - a$. Similarly for $(a, b]$.

**1.6.5**

Consider the transformation $Tx = \alpha x + \beta$ from the real line onto itself, where $\alpha, \beta$ are real numbers and $\alpha \neq 0$. It maps sets $E$ onto sets $T(E)$. Denote by $\mu$ ($\mu^*$) the Lebesgue measure (outer measure) on the real line. Prove

(a) For any set $E$, $\mu^*(T(E)) = |\alpha|\mu^*(E)$.

(b) $E$ is Lebesgue-measurable if and only if $T(E)$ is Lebesgue-measureable.

(c) If $E$ is Lebesgue-measurable, then $\mu(T(E)) = |\alpha|\mu(E)$.

*Solution.* Let us start with a couple of simple observations:

- $T$ is bijective, with inverse given by

$$T^{-1}(x) = \frac{x - \beta}{\alpha}.$$

- Suppose $I = (a, b)$. Then

$$T(I) = (\alpha a + \beta, \alpha b + \beta)$$

if $\alpha > 0$, and

$$T(I) = (\beta b + \beta, \alpha a + \beta)$$

if $\alpha < 0$. Either way,

$$\mu^*[T(I)] = |\alpha|(b - a) = |\alpha|\mu^*(I),$$

where we have used one of the results of the previous exercise. Similarly, $T^{-1}(I)$ is an open interval and

$$\mu^*[T^{-1}(I)] = |\alpha|^{-1}\mu^*(I).$$

Of course, the latter two identities still hold if $I = \varnothing$. Hence they hold for all $I \in \mathscr{K}$.

Also, let us use the notation

$$\Lambda(E) = \left\{ \sum_{n=1}^{\infty} \lambda(I_n);\ I_n \in \mathscr{K},\ \bigcup_{n=1}^{\infty} I_n \supset E \right\}$$

for all $E \subset \mathbb{R}$.

(a) Suppose $(I_n)$ is a sequence in $\mathscr{K}$ (i.e. a sequence of open intervals) and $E \subset \bigcup_n I_n$. Then $T(I_n) \in \mathscr{K}$ for every $n$,

$$T(E) \subset T\left( \bigcup_n I_n \right) = \bigcup_n T(I_n),$$

10

and

$$\sum_n \lambda[T(I_n)] = |\alpha| \sum_n \lambda(I_n).$$

Thus, if $s \in \Lambda(E)$, then $|\alpha| s \in \Lambda[T(E)]$. It follows that

$$\mu^*[T(E)] = \inf \Lambda[T(E)] \le |\alpha| \inf \Lambda(E) = |\alpha| \mu^*(E).$$

Conversely, suppose $(J_n)$ is a sequence in $\mathscr{K}$ and $T(E) \subset \bigcup_n J_n$. Then $T^{-1}(J_n) \in \mathscr{K}$ for all $n$,

$$E = T^{-1}[T(E)] \subset T^{-1}\left(\bigcup_n J_n\right) = \bigcup_n T^{-1}(J_n),$$

and

$$\sum_n \lambda[T^{-1}(J_n)] = |\alpha|^{-1} \sum_n \lambda(J_n).$$

Hence, by the same logic as above, we find that $\mu^*(E) \le |\alpha|^{-1} \mu^*[T(E)]$, and it follows that

$$\mu^*[T(E)] = |\alpha| \mu^*(E).$$

(b) Note that if $f : X \to Y$ is a bijective function (between arbitrary sets $X, Y$), then

$$f^{-1}[f(A)] = A,$$
$$f(A \cup B) = f(A) \cup f(B),$$
$$f(A - B) = f(A) - f(B),$$
$$f[f^{-1}(C)] = C,$$

for all $A, B \subset X$ and $C \subset Y$.

Suppose that $E$ is measureable:

$$\mu^*(A) = \mu^*(A \cap E) + \mu^*(A - E)$$

for all $A \subset \mathbb{R}$. Then, for all $B \subset \mathbb{R}$, we have

$$
\begin{aligned}
\mu^*[B \cap T(E)] &+ \mu^*[B - T(E)] \\
&= \mu^*[T(T^{-1}(B) \cap E)] + \mu^*[T(T^{-1}(B) - E)] \\
&= |\alpha| \mu^*[T^{-1}(B) \cap E] + |\alpha| \mu^*[T^{-1}(B) - E] \\
&= |\alpha| \mu^*[T^{-1}(B)] \\
&= \mu^*(B),
\end{aligned}
$$

so that $T(E)$ is measurable.

Conversely, suppose that $T(E)$ is measurable. Then, for all $A \subset \mathbb{R}$,

$$
\begin{aligned}
\mu^*(A \cap E) &+ \mu^*(A - E) \\
&= \mu^*[T^{-1}(T(A) \cap T(E))] + \mu^*[T^{-1}(T(A) - T(E))] \\
&= |\alpha|^{-1} \mu^*[T(A) \cap T(E)] + |\alpha|^{-1} \mu^*[T(A) - T(E)] \\
&= |\alpha|^{-1} \mu^*[T(A)] \\
&= \mu^*(A),
\end{aligned}
$$

so that $E$ is measurable.

(c) This is immediate given (a), (b), and the definition of the Lebesgue-measure. First, $T(E)$ is Lebesgue-measurable by (b). Next, $\mu(E) = \mu^*(E)$ and $\mu[T(E)] = \mu^*[T(E)]$ since $\mu$ is simply the restriction of $\mu^*$ to the measurable sets. Finally, $\mu^*[T(E)] = |\alpha|\mu^*(E)$ by (a).

# Section 1.7 – Metric Spaces

## Problems

### 1.7.5

Let $A$ be any set and $x, y$ any two points. If $\rho(x, A) > \alpha$, $\rho(y, A) < \beta$ and $\alpha > \beta$, then $\rho(x, y) > \alpha - \beta$.

*Solution.* Since $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$ for all $z \in A$ (by the triangle inequality), we have

$$\alpha < \rho(x, A) \leq \rho(x, y) + \rho(y, A) < \rho(x, y) + \beta,$$

whereby the result follows.

# Section 1.8 – Metric Outer Measure

## Problems

### 1.8.1

Prove the converse of Corollary 1.8.3 – that is, if $\mu^*$ is an outer measure and if every open set is measurable, then $\mu^*$ is a metric outer measure.

*Solution.* Suppose $A$ and $B$ are sets such that $\rho(A, B) > 0$. Let $d = \rho(A, B)$, and define $E = \bigcup_{x \in A} B(x, d)$. Then $E$ is open (hence measurable), $A \subset E$, and $B \subset E^c$. Thus

$$\mu^*(A \cup B) = \mu^*((A \cup B) \cap E) + \mu^*((A \cup B) \setminus E) = \mu^*(A) + \mu^*(B)$$

by (1.3.3). This shows that $\mu^*$ is a metric outer measure.

### 1.8.3

Let $(X, \rho)$ be a metric space and let $\{x_n\}$ be a sequence of points in $X$. Define $\mu^*(E)$ to be the number of points $x_n$ that belong to $E$. Prove that $\mu^*$ is a metric outer measure.

*Solution.* To show that $\mu^*$ is an outer measure one can argue similarly to Problem 1.3.1; hence we will only show that it satisfies property (vi).

Suppose $A$ and $B$ are subsets of $X$ such that $\rho(A, B) > 0$. Then $A$ and $B$ are disjoint, so the number of points of $\{x_n\}$ that lie in $A \cup B$ is simply the sum of the number that lie in $A$ and the number that lie in $B$. That is,

$$\mu^*(A \cup B) = \mu^*(A) + \mu^*(B).$$

**1.8.4**

Let $(X, \rho)$ be a metric space. Define $\mu^*(E) = 1$ if $E \neq \varnothing$ and $\mu^*(\varnothing) = 0$. Is $\mu^*$ a metric outer measure?

*Solution.* It was shown in Problem 1.3.2 that $\mu^*$ is an outer measure, but it is *not* a *metric* outer measure (unless $|X| < 2$). Indeed, if $X$ contains two distinct elements $x$ and $y$, take $A = \{x\}$ and $B = \{y\}$. Then $\rho(A, B) = \rho(x, y) > 0$, but

$$\mu^*(A \cup B) = 1 \neq 2 = \mu^*(A) + \mu^*(B).$$

# Section 1.9 – Construction of Metric Outer Measures

## Problems

### 1.9.7

If a set $F$ in $R^n$ is Lebesgue-measurable and if $\mu(F) < \infty$, then for any $\epsilon > 0$ there exists an open set $E$ such that $E \supset F$ and $\mu(E) < \mu(F) + \epsilon$.

*Solution.* Let $\mathscr{K}$, $\lambda$, $\mu^*$, and $\mu$ be as in Section 1.6. By the definition of the Lebesgue measure via the Lebesgue outer measure, there exist $\{E_n\} \subset \mathscr{K}$ such that $F \subset E := \bigcup_n E_n$ and

$$\sum_n \lambda(E_n) < \mu(F) + \epsilon.$$

Since every element of $\mathscr{K}$ is open, $E$ is open; thus $E$ is Lebesgue-measurable by Problem 1.9.3. It follows that

$$\mu(E) \leq \sum_n \lambda(E_n),$$

hence $\mu(E) < \mu(F) + \epsilon$.

### 1.9.8

If a set $F$ in $R^n$ is Lebesgue-measurable, then there exists a Borel set $E$ such that $E \supset F$ and $\mu(E - F) = 0$. [Hint: Consider first the case where $\mu(F) < \infty$.]

*Solution.* Assume for now that $\mu(F) < \infty$. By the previous problem, for every $n \geq 1$ there exists an open set $E_n$ such that $E_n \supset F$ and $\mu(E_n) < \mu(F) + 1/n$. Let $E = \underline{\lim}_n E_n$; then $E$ is a Borel set, and measurable by Problem 1.9.3. Moreover, $\mu(F) \geq \mu(E)$ since $F \supset E$, and

$$\mu(E) \leq \underline{\lim_n} \mu(E_n) \leq \mu(F)$$

by Theorem 1.2.2. It follows that

$$\mu(E - F) = \mu(E) - \mu(F) = 0.$$

Now remove the assumption that $\mu(F) < \infty$. Fix a sequence $A_1, A_2, \ldots$ of Borel sets such that $X = \bigcup_n A_n$ and $\mu(A_n) < \infty$. For example, by Problem 1.9.6, we can take the sets $A_n$ to be a suitable family of open intervals of increasing size. For each $n$, let $F_n = F \cap \bigcup_{i=1}^n A_i$, so that $\mu(F_n) < \infty$. By the previous argument, there exists a Borel set $G_n$ such that $G_n \supset F_n$ and $\mu(G_n - F_n) = 0$. Define $G = \underline{\lim}_n G_n$; then $G_n$ is a Borel set (hence measurable). Since

$$\underline{\lim_n}(G_n - F_n) = \underline{\lim_n} G_n - \overline{\lim_n} F_n = G - F,$$

we have

$$\mu(G - F) \le \underline{\lim_n} \mu(G_n - F_n) = 0.$$

# Section 1.10 – Signed Measures

## Problems

### 1.10.1

If $X = \bigcup_{n=1}^\infty A_n$ and $\mu$ is a signed measure with $|\mu(A_n)| < \infty$ for all $n$, then $\mu^+$ and $\mu^-$ are $\sigma$-finite. (Hence, by definition, $\mu$ is $\sigma$-finite.)

*Solution.* For all $n$ we have $|\mu^+(A_n) - \mu^-(A_n)| = |\mu(A_n)| < \infty$, hence $\mu^+(A_n) < \infty$ and $\mu^-(A_n) < \infty$. It follows that $\mu^+$ and $\mu^-$ are $\sigma$-finite.

### 1.10.2

Let $\mu$ be a signed measure and let $\{E_n\}$ be a monotone sequence of measurable sets [with $|\mu(E_1)| < \infty$ if $\{E_n\}$ decreasing]. Then

$$\mu\left(\lim_n E_n\right) = \lim_n \mu(E_n).$$

*Solution.* Note that

$$\mu^\pm\left(\lim_n E_n\right) = \lim_n \mu^\pm(E_n)$$

by Theorem 1.2.1. Hence

$$\mu\left(\lim_n E_n\right) = \mu^+\left(\lim_n E_n\right) - \mu^-\left(\lim_n E_n\right) = \lim_n \mu^+(E_n) - \lim_n \mu^-(E_n),$$

while

$$\lim_n \mu(E_n) = \lim_n \left[\mu^+(E_n) - \mu^-(E_n)\right].$$

If $\mu^\pm(\lim_n E_n)$ are both finite, then the sequences $\mu^\pm(E_n)$ are convergent, and the result follows immediately. If $\mu^+(\lim_n E_n) = \infty$, then $\mu^-$ must be a finite measure. It follows that the sequence $\mu^-(E_n)$ is bounded by some $K > 0$, hence

$$\mu\left(\lim_n E_n\right) \ge \infty - K = \infty$$

and
$$\lim_n \mu(E_n) \geq \lim_n \left[ \mu^+(E_n) - K \right] = \mu^+ \left( \lim_n E_n \right) - K = \infty,$$

so that $\mu \left( \lim_n E_n \right) = \lim_n \mu(E_n)$ again. The argument when $\mu^- \left( \lim_n E_n \right) = \infty$ is similar.

### 1.10.3

Give an example of a signed measure for which the Hahn decomposition is not unique.

*Solution.* Let $X = \{0, 1\}$, let $\mathcal{A}$ be the power set of $X$, and let $\mu(E) = \sum_{x \in E} x$. Then $\mu$ is a signed measure; in fact it is a measure. One Hahn decomposition is given by $A = X$ and $B = \varnothing$, another by $A = \{1\}$ and $B = \{0\}$.

### 1.10.4

If $X = A_1 \cup B_1$, $X = A_2 \cup B_2$ are two Hahn decompositions of a signed measure $\mu$, then for any measurable set $E$,

$$\mu(E \cap A_1) = \mu(E \cap A_2), \quad \mu(E \cap B_1) = \mu(E \cap B_2).$$

*Solution.* Consider $\mu(E \cap A_1 \cap B_2)$. On the one hand it is nonnegative, since $A_1$ is positive with respect to $\mu$, on the other it is nonpositive, since $B_2$ is negative. Hence $\mu(E \cap A_1 \cap B_2) = 0$, and similarly $\mu(E \cap A_2 \cap B_1) = 0$. It follows that

$$\begin{aligned}
\mu(E \cap A_1) &= \mu(E \cap A_1 \cap A_2) + \mu(E \cap A_1 \cap B_2) \\
&= \mu(E \cap A_2 \cap A_1) + \mu(E \cap A_2 \cap B_1) \\
&= \mu(E \cap A_2).
\end{aligned}$$

The argument that $\mu(E \cap B_1) = \mu(E \cap B_2)$ is similar.

### 1.10.5

A *complex measure* is, by definition, a finite complex-valued set functions satisfying the properties (i), (iii), (iv) of Definition 1.2.1. Prove that $\mu$ is a complex measure with domain $\mathcal{A}$ if and only if their exist finite measures $\mu_1, \mu_2, \mu_3, \mu_4$ with the same domain $\mathcal{A}$ such that

$$\mu = \mu_1 - \mu_2 + i\mu_3 - i\mu_4.$$

*Solution.* Suppose $\mu$ is a complex measure. Define $\alpha = \operatorname{Re} \mu$ and $\beta = \operatorname{Im} \mu$. Then $\alpha$ and $\beta$ are easily seen to be finite signed measures, so they each have a Jordan decomposition with $\alpha^\pm$ and $\beta^\pm$ finite, and

$$\mu = \alpha^+ - \alpha^- + i\beta^+ - i\beta^-.$$

The converse implication is pretty much immediate.

# Chapter 2 – Integration

## Section 2.1 – Definition of Measurable Functions

### Problems

**2.1.6**

The *characteristic function* of a set $E$ is the function $\chi_E$ defined by

$$\chi_E(x) = \begin{cases} 1, & \text{if } x \in E, \\ 0, & \text{if } x \notin E. \end{cases}$$

Prove that the set $E$ is measurable if and only if the function $\chi_E$ is measurable.

*Solution.* Suppose $E \in \mathcal{A}$. For all $c \in \mathbb{R}$,

$$\chi_E^{-1}\{(-\infty, c)\} = \{x \in X; \chi_E(x) < c\} = \begin{cases} \varnothing & (c \leq 0), \\ E^c & (0 < c \leq 1), \\ X & (c > 1), \end{cases}$$

so that $\chi_E^{-1}\{(-\infty, c)\} \in \mathcal{A}$. By Theorem 2.1.1, $\chi_E$ is measurable.

Conversely, suppose $\chi_E$ is measurable. Then $E$ is measurable, since

$$E = X - E^c = \chi^{-1}\{(-\infty, 2)\} - \chi^{-1}\{(-\infty, 1)\}.$$

**2.1.9**

If $f$ is measurable, then $|f|$ and $|f|^2$ are measurable.

*Solution.* If $c \leq 0$, then

$$(|f|)^{-1}\{(-\infty, c)\} = (|f|^2)^{-1}\{(-\infty, c)\} = \emptyset \in \mathcal{A},$$

since $|f|$ and $|f|^2$ are nonnegative functions.

Let $c > 0$. Then

$$(|f|)^{-1}\{(-\infty, c)\} = \{x \in X; -c < f(x) < c\} = f^{-1}\{(-c, c)\}.$$

The set $(-c, c)$ is open, hence $f^{-1}\{(-c, c)\} \in \mathcal{A}$ by the measurability of $f$. Similarly,

$$(|f|^2)^{-1}\{(-\infty, c)\} = f^{-1}\{(-\sqrt{c}, \sqrt{c})\} \in \mathcal{A}.$$

Finally,

$$(|f|)^{-1}\{+\infty\} = (|f|^2)^{-1}\{+\infty\} = f^{-1}\{+\infty\} \cup f^{-1}\{-\infty\} \in \mathcal{A}$$

by the measurability of $f$, and

$$(|f|)^{-1}\{-\infty\} = (|f|^2)^{-1}\{-\infty\} = \emptyset \in \mathcal{A}$$

since $|f|$ and $|f|^2$ are nonnegative. Thus, both $|f|$ and $|f|^2$ are measurable by Theorem 2.1.1.

### 2.1.10

A monotone function defined on the real line is Lebesgue-measurable.

*Solution.* Let $f$ be a monotone increasing extended real-valued function on $\mathbb{R}$;

$$(\forall x, y \in \mathbb{R}): \quad x < y \implies f(x) \leq f(y).$$

Given any $c \in \mathbb{R}$, let

$$\xi_c = \inf\{x \in X; f(x) \geq c\}.$$

We need to consider two cases: $f(\xi_c) < c$ and $f(\xi_c) \geq c$. In the former case, $f(x) < c$ for all $x \leq \xi_c$ and $f(x) \geq c$ for all $x > \xi_c$ (by monotonicity). Hence

$$f^{-1}\{(-\infty, c)\} = (-\infty, \xi_c].$$

This is a Borel set, hence also a Lebesgue set (see Problem 1.9.3). In the latter case, $f(x) < c$ for all $x < \xi_c$ and $f(x) \geq c$ for all $x \geq \xi_c$, so that

$$f^{-1}\{(-\infty, c)\} = (-\infty, \xi_c),$$

which is Lebesgue-measurable. Since $c$ was arbitrary, we conclude that $f$ is measurable, by Theorem 2.1.1.

The proof for $f$ monotone decreasing is similar.

# Section 2.2 – Operations on Measurable Functions

## Problems

### 2.2.2

Let $g(u_1, \ldots, u_k)$ be a continuous function in $\mathbb{R}^k$, and let $\varphi_1, \ldots, \varphi_k$ be measurable functions. Prove that the composite function $h(x) = g[\varphi_1(x), \ldots, \varphi_k(x)]$ is a measurable function. Note that as a special case we may conclude that

$$\max(\varphi, \ldots, \varphi_n) \quad \text{and} \quad \min(\varphi, \ldots, \varphi_n)$$

are measurable functions.

*Solution.* We will use the following fact, which may be proven in a course in topology:

> $\mathbb{R}^k$ has a countable basis of product open subsets. Hence, if $U$ is an open subset of $\mathbb{R}^k$, then there are open subsets $U_{ni} \subset \mathbb{R}$ for $n = 1, 2, \ldots$ and $i = 1, \ldots, k$ such that
> $$U = \bigcup_{n=1}^{\infty} (U_{n1} \times \cdots \times U_{nk}).$$

We are assuming that $g$ is real-valued, likewise for the functions $\varphi_i$. Let $c \in \mathbb{R}$. Note that $g^{-1}\{(-\infty, c)\}$ is open by continuity of $g$. Thus

$$g^{-1}\{(-\infty, c)\} = \bigcup_{n=1}^{\infty} (U_{n1} \times \cdots \times U_{nk})$$

for some open subsets $U_{ni} \subset \mathbb{R}$. Hence

$$
\begin{aligned}
h^{-1}\{(-\infty, c)\} &= \{x \in X; g(\varphi_1(x), \ldots, \varphi_k(x)) \le c\} \\
&= \{x \in X; (\varphi_1(x), \ldots, \varphi_k(x)) \in g^{-1}\{(-\infty, c)\}\} \\
&= \bigcup_{n=1}^{\infty} \{x \in X; (\varphi_1(x), \ldots, \varphi_k(x)) \in U_{n1} \times \cdots \times U_{nk}\} \\
&= \bigcup_{n=1}^{\infty} \bigcap_{i=1}^{k} \{x \in X; \varphi_i(x) \in U_{ni}\} \\
&= \bigcup_{n=1}^{\infty} \bigcap_{i=1}^{k} \varphi_i^{-1}(U_{ni}).
\end{aligned}
$$

The sets $\varphi_i(U_{ni})$ are measurable since the functions $\varphi_i$ are measurable. It follows that $h^{-1}\{(-\infty, c)\}$ is measurable, and thus that $h$ is measurable, by Theorem 2.1.1.

To apply the above to the max and min functions $\mathbb{R}^k \to \mathbb{R}$ we must show that they are continuous. Let $a < b$ and note that

$$
\begin{aligned}
\max{}^{-1}\{(a, b)\} &= \{(x_1, \ldots, x_k) \in \mathbb{R}^k; x_i > a \text{ for some } i\} \\
&\quad \cap \{(x_1, \ldots, x_k) \in \mathbb{R}^k; x_i < b \text{ for all } i\}.
\end{aligned}
$$

Both sets in the above binary intersection are easily seen to be open by considering $\epsilon$-neighborhoods about their points. It follows that $\max^{-1}(U)$ is open for all open subsets $U \in \mathbb{R}^k$, since every such $U$ can be written as a countable union of open intervals. Thus max is continuous, and one similarly shows that min is continuous.

**2.2.3**

Let $f(x)$ be a measurable function and define

$$g(x) = \begin{cases} \frac{1}{f(x)}, & \text{if } f(x) \neq 0, \\ 0, & \text{if } f(x) = 0. \end{cases}$$

Prove that $g$ is measurable.

*Solution.* For $c < 0$,

$$g^{-1}\{(-\infty, c)\} = \{x; 1/f(x) < c\} = \{x; 1/c < f < 0\} = f^{-1}\{(1/c, 0)\},$$

which is measurable by the measurability of $f$. Next,

$$g^{-1}\{(-\infty, 0)\} = \{x; 1/f(x) < 0\} = \{x; f(x) < 0\} = f^{-1}\{(-\infty, 0)\},$$

also measurable. Note that if we take the natural convention (unfortunately not addressed in the text) that $x/(\pm\infty) = 0$ for all $x \in \mathbb{R}$, then

$$g^{-1}(\{0\}) = \{x; f(x) = 0\} \cup \{x; f(x) = \pm\infty\} = f^{-1}(\{0\}) \cup f^{-1}(\{\pm\infty\}).$$

Hence, for $c > 0$,

$$\begin{aligned} g^{-1}\{(-\infty, c)\} &= g^{-1}\{(-\infty, 0)\} \cup g^{-1}(\{0\}) \cup g^{-1}\{(0, c)\} \\ &= f^{-1}\{(-\infty, 0)\} \cup f^{-1}(\{0\}) \cup f^{-1}(\{\pm\infty\}) \cup f^{-1}\{(1/c, \infty)\} \\ &= f^{-1}\{(-\infty, 0]\} \cup f^{-1}(\{\pm\infty\}) \cup f^{-1}\{(1/c, \infty)\}, \end{aligned}$$

which is measurable by the measurability of $f$ (see Problem 2.1.4). Finally, $g^{-1}(\{\pm\infty\}) = \emptyset$, and it follows by Theorem 2.1.1 that $g$ is measurable.

# Section 2.3 – Egoroff's Theorem

## Problems

### 2.3.2

Let $\{f_n\}$ be a sequence of measurable functions in a finite measure space $X$. Suppose that for almost every $x$, $\{f_n(x)\}$ is a bounded set. Then for any $\epsilon > 0$ there exist a positive number $c$ and a measurable set $E$ with $\mu(X - E) < \epsilon$, such that $|f_n(x)| \leq c$ for all $x \in E$, $n = 1, 2, \ldots$.

*Solution.* The definition we have for 'bounded set' applies to metric spaces, and it does not make much sense here since the functions $f_n$ may be extended real-valued. Hence we will assume that '$\{f_n(x)\}$ is a bounded set' means that $\sup_n |f_n(x)| < \infty$.

Let $g = \sup_n |f_n|$, and note that $g$ is measurable by Problem 2.1.9 and Theorem 2.2.3. Let $F = \{x; g(x) < \infty\}$. Notice that $g(x) < \infty$ if and only if $\{f_n(x)\}$ is bounded. Hence $\mu(X - F) = 0$.

For $k = 1, 2, \ldots$, define $F_k = \{x; g(x) \leq k\}$. Then $F_1 \subset F_2 \subset \cdots$ and $\lim_k F_k = \bigcup_{k=1}^{\infty} F_k = F$. By Theorem 1.2.1(iv),

$$\lim_k \mu(X - F_k) = \mu(X - F) = 0.$$

Given any $\epsilon > 0$, there exists a positive integer $K$ such that $\mu(X - F_k) < \epsilon$ for all $k \geq K$. In particular $\mu(X - F_K) < \epsilon$, and $g(x) \leq K$ for all $x \in F_K$, which means that $|f_n(x)| \leq K$ for all $x \in F_K$.

# Section 2.4 – Convergence in Measure

## Problems

### 2.4.3

Prove the following result (which immediately yields another proof of Corollary 2.4.2): Let $f_n$ $(n = 1, 2, \ldots)$ and $f$ be a.e. real-valued measurable functions in a finite measure space. For any $\epsilon > 0$, $n \geq 1$, let

$$E_n(\epsilon) = \{x; |f_n(x) - f(x)| \geq \epsilon\}.$$

Then $\{f_n\}$ converges a.e. to $f$ if and only if

$$\lim_{n \to \infty} \mu \left[ \bigcup_{m=n}^{\infty} E_m(\epsilon) \right] = 0 \quad \text{for any } \epsilon > 0. \tag{2.4.2}$$

[*Hint:* Let $F = \{x; \{f_n(x)\}$ is not convergent to $f(x)\}$. Then $F = \bigcup_{k=1}^{\infty} \overline{\lim}_n E_n(1/k)$. Show that $\mu(F) = 0$ if and only if (2.4.2) holds.]

*Solution.* Define

$$F = \bigcup_{k=1}^{\infty} \overline{\lim_n} \, E_n \left( \frac{1}{k} \right) = \bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m \left( \frac{1}{k} \right).$$

Note that

$$x \in F \iff \exists k, \forall n, \exists m \geq n, \ |f_m(x) - f(x)| \geq \frac{1}{k}$$

$$\iff \neg \left( \forall k, \exists n, \forall m \geq n, \ |f_m(x) - f(x)| < \frac{1}{k} \right)$$

$$\iff f_n(x) \not\to f(x),$$

so that

$$F = \{x; f_n(x) \not\to f(x)\}.$$

Suppose (2.4.2) holds. Fix $\delta > 0$. For every positive integer $k$, there exists a positive integer $n_k$ such that $n \geq n_k$ implies

$$\mu \left[ \bigcup_{m=n}^{\infty} E_m \left( \frac{1}{k} \right) \right] < \frac{\delta}{2^k}.$$

By subadditivity and monotonicity,

$$\mu(F) = \mu\left[\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m\left(\frac{1}{k}\right)\right] \leq \sum_{k=1}^{\infty} \mu\left[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m\left(\frac{1}{k}\right)\right]$$

$$\leq \sum_{k=1}^{\infty} \mu\left[\bigcup_{m=n_k}^{\infty} E_m\left(\frac{1}{k}\right)\right] < \sum_{k=1}^{\infty} \frac{\delta}{2^k} = \delta.$$

Since $\delta$ was arbitrary, $\mu(F) = 0$, and it follows that $f_n \to f$ a.e.

Conversely, suppose $f_n \to f$ a.e., so that $\mu(F) = 0$. By monotonicity and Theorem 1.2.2,

$$0 = \mu(F) = \mu\left[\bigcup_{k=1}^{\infty} \overline{\lim_n} E_n\left(\frac{1}{k}\right)\right] \geq \mu\left[\overline{\lim_n} E_n\left(\frac{1}{l}\right)\right] \geq \overline{\lim_n} \mu\left[E_n\left(\frac{1}{l}\right)\right]$$

for all positive integers $l$. But of course $\overline{\lim}_n \mu\left[E_n\left(1/l\right)\right] \geq \underline{\lim}_n \mu\left[E_n\left(1/l\right)\right] \geq 0$ since $\mu$ is nonnegative, so $\lim_n \mu\left[E_n\left(1/l\right)\right]$ exists and is equal to zero. Note that the sets $\bigcup_{m=n}^{\infty} E_m(1/l)$ are decreasing, so their limit as $n \to \infty$ exists. Hence we can apply Corollary 1.2.3 and monotonicity to find that

$$\lim_n \mu\left[\bigcup_{m=n}^{\infty} E_m\left(\frac{1}{l}\right)\right] = \mu\left[\lim_n \bigcup_{m=n}^{\infty} E_m\left(\frac{1}{l}\right)\right] = \mu\left[\bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m\left(\frac{1}{l}\right)\right]$$

$$\leq \mu\left[\bigcup_{k=1}^{\infty} \bigcap_{n=1}^{\infty} \bigcup_{m=n}^{\infty} E_m\left(\frac{1}{k}\right)\right] = \mu(F) = 0.$$

Finally, given $\epsilon > 0$, note that

$$E_n(\epsilon) \subset E_n\left(\frac{1}{\lceil 1/\epsilon \rceil}\right).$$

Hence

$$\lim_n \mu\left[\bigcup_{m=n}^{\infty} E_m(\epsilon)\right] \leq \lim_n \mu\left[\bigcup_{m=n}^{\infty} E_m\left(\frac{1}{\lceil 1/\epsilon \rceil}\right)\right] \leq 0$$

by monotonicity, and (2.4.2) follows.

**2.4.4**

Let $X$ be the set of all positive integers, $\mathcal{A}$ the class of all subsets of $X$, and $\mu(E)$ (for any $E \in \mathcal{A}$) the number of points in $E$. Prove that in this measure space, convergence in measure is equivalent to uniform convergence.

*Solution.* Uniform convergence always implies convergence in measure. Conversely, suppose $(f_n)$ converges in measure to $f$. Given any $\epsilon > 0$, there exists a positive integer $N$ such that $n \geq N$ implies

$$\mu\left[\{x;\, |f_n(x) - f(x)| \geq \epsilon\}\right] < 1.$$

That is, for $n \geq N$ the set $\{x;\, |f_n(x) - f(x)| \geq \epsilon\}$ is empty, which in particular means that $\sup_x |f_n(x) - f(x)| \leq \epsilon$. It follows that $f_n \to f$ uniformly.

21

# Section 2.5 – Integrals of Simple Functions

## Problems

### 2.5.2

An integrable simple function $f$ is equal a.e. to zero if and only if $\int_E f d\mu = 0$ for any measurable set $E$.

*Solution.* Let $f$ be an integrable simple function. Then $f$ can be written in the form

$$f = \sum_{i=1}^{n} \alpha_i \chi_{E_i},$$

for mutually disjoint sets $E_1, \ldots E_n$, with all $\alpha_i \neq 0$, and all $\mu(E_i) < \infty$.

Suppose $f = 0$ a.e., and let $E$ be any measurable set. By Theorem 2.5.1(b) and (g),

$$0 \leq \int_E f d\mu \leq \int f d\mu = \sum_{i=1}^{n} \alpha_i \mu(E_i).$$

But $\mu(E_i) = 0$ since $f = 0$ a.e., so $\int_E f d\mu = 0$.

Conversely, suppose $\int_E f d\mu = 0$ for all measurable sets $E$. Then

$$\alpha_i \mu(E_i) = \int_{E_i} f d\mu = 0,$$

so that $\mu(E_i) = 0$, for all $i \in \{1, \ldots, n\}$. It follows that $f = 0$ a.e.

# Section 2.6 – Definition of the Integral

## Problems

### 2.6.3

Let $f$ be a measurable function. Prove that $f$ is integrable if and only if $f^+$ and $f^-$ are integrable, or if and only if $|f|$ is integrable.

*Solution.* Let $f$ be measurable. We must prove the equivalence of the following statements:

(i)  $f$ is integrable.

(ii)  $f^+$ and $f^-$ are integrable.

(iii)  $|f|$ is integrable.

We will first show that (iii) $\implies$ (ii), then that (ii) $\implies$ (i), and finally that (i) $\implies$ (iii).

Suppose that $|f|$ is integrable. Let $E = \{x;\ f(x) \geq 0\} = f^{-1}[0, \infty)$, and note that $E$ is measurable since $f$ is. There exists a Cauchy in the mean sequence $(g_n)$ of integrable simple functions converging to $|f|$ a.e., and the sequence $(\chi_E g_n)$

is easily seen to satisfy the corresponding properties with respect to $f^+$. Since $f^+$ is measurable by Problem 2.1.8, this implies that it is integrable. The proof that $f^-$ is integrable is similar.

Next, suppose that $f^+$ and $f^-$ are integrable. Then there exist Cauchy in the mean sequences $(g_n)$ and $(h_n)$ of integrable simple functions converging a.e. to $f^+$ and $f^-$, respectively. Define a new sequence $(f_n)$ of integrable simple functions by $f_n = g_n - h_n$. Then $(f_n)$ is Cachy in the mean, since

$$|f_n - f_m| = |g_n - h_n - g_m + h_m| \leq |g_n - g_m| + |h_n - h_m|.$$

It also converges to $f$ a.e. since

$$|f_n - f| = |g_n - h_n - f^+ + f^-| \leq |g_n - f^+| + |h_n - f^-|.$$

It follows that $f$ is integrable.

Finally, assume that $f$ is integrable. There is a Cauchy in the mean sequence $(f_n)$ of integrable simple functions converging to $f$ a.e. The sequence $(|f_n|)$ consists of integrable simple functions. It is Cauchy in the mean since

$$||f_n| - |f_m|| \leq |f_n - f_m|,$$

and it converges to $|f|$ a.e. since

$$||f_n| - |f|| \leq |f_n - f|.$$

Since $|f|$ is measurable by Problem 2.1.9, it follows that $|f|$ is integrable.

### 2.6.4

Let $X$ be the measure space described in Problem 2.4.4. Then $f$ is integrable if and only if the series $\sum_{n=1}^{\infty} |f(n)|$ is convergent. If $f$ is integrable, then

$$\int f \, d\mu = \sum_{n=1}^{\infty} f(n).$$

*Solution.* Suppose $f$ is integrable. Then there is a Cauchy in the mean sequence $(f_n)$ of integrable simple functions converging to $f$ a.e. We saw in the previous problem that this implies that $|f|$ is integrable, and that $(|f_n|)$ is a Cauchy in the mean sequence of integrable simple functions converging to $|f|$ a.e. Note that in this particular space convergence a.e. is the same as convergence everywhere (since the only subset with measure zero is $\emptyset$).

By Theorem 2.5.1(h),

$$\int |f_n| \, d\mu = \sum_{i=1}^{\infty} \int_{\{i\}} |f_n| \, d\mu = \sum_{i=1}^{\infty} |f_n(i)|.$$

Hence, in particular,

$$\int |f| \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{\infty} |f_n(i)|.$$

23

Given any positive integer $m$, there exists $n'$ such that

$$|f(i) - f_{n'}(i)| < 1/m \quad (i = 1, 2, \ldots, m)$$

(since $f_n \to f$) and

$$\left| \sum_{i=1}^{\infty} |f_{n'}(i)| - \int |f| \, d\mu \right| < 1$$

(since $\sum_i |f_n(i)| \to \int |f| \, d\mu$). Thus

$$\sum_{i=1}^{m} |f(i)| \leq \sum_{i=1}^{m} |f(i) - f_{n'}(i)| + \sum_{i=1}^{m} |f_{n'}(i)| < 1 + \sum_{i=1}^{\infty} |f_{n'}(i)| < 2 + \int |f| \, d\mu,$$

and it follows that the series $\sum_{i=1}^{\infty} |f(i)|$ converges (to a finite number).

Conversely, assume that the series $\sum_{i=1}^{\infty} |f(i)|$ converges. Define a sequence of integrable simple functions $(g_n)$ by

$$g_n = \sum_{i=1}^{n} f(i) \chi_{\{i\}}.$$

It is clear that $g_n \to f$ everywhere. Moreover, if $m > n$, then

$$\int |g_m - g_n| \, d\mu = \int \left| \sum_{i=n+1}^{m} f(i) \chi_{\{i\}} \right| \, d\mu = \sum_{n+1}^{m} |f(i)| \leq \sum_{n+1}^{\infty} |f(i)|.$$

The right-hand side goes to zero as $n \to \infty$ since $\sum_{i=1}^{\infty} |f(i)|$ is convergent, which means that $\int |g_m - g_n| \, d\mu \to 0$ as $n, m \to \infty$; i.e., $(g_n)$ is Cauchy in the mean. It follows that $f$ is integrable, with

$$\int f \, d\mu = \lim_{n \to \infty} \int g_n \, d\mu = \lim_{n \to \infty} \sum_{i=1}^{n} f(i) = \sum_{i=1}^{\infty} f(i).$$

# Section 2.7 – Elementary Properties of Integrals

## Problems

### 2.7.3

Let $f$ be an integrable function. Prove: (a) if $\int_E f \, d\mu \geq 0$ for all measurable sets $E$, then $f \geq 0$ a.e.; (b) if $\mu(X) < \infty$ and if $\int_E f \, d\mu \leq \mu(E)$ for all measurable sets $E$, then $f \leq 1$ a.e.

*Solution.*

(a) Let $F = \{x; f(x) < 0\}$. Then $-f$ is positive on $F$, so

$$-\int_F f \, d\mu = \int_F (-f) \, d\mu \geq 0$$

24

by Theorem 2.7.1(b). But

$$\int_F f d\mu \geq 0$$

by our hypothesis on $f$, so

$$\int_F f d\mu = 0,$$

and it follows by Theorem 2.7.5 that $\mu(F) = 0$. That is, $f \geq 0$ a.e.

(b) Let $G = \{x : f(x) > 1\}$. Then $f - 1$ is positive on $G$, so

$$\int_G (f - 1) d\mu \geq 0$$

by Theorem 2.7.1(b). Note that $\chi_G$ is integrable since $\mu(G) \leq \mu(X) < \infty$. Hence we also have

$$\int_G (f - 1) d\mu = \int_G f d\mu - \int_G d\mu = \int_G f d\mu - \mu(G) \leq 0,$$

with the inequality following from our hypothesis on $f$. Thus

$$\int_G (f - 1) d\mu = 0,$$

and Theorem 2.7.5 yields $\mu(G) = 0$. That is, $f \leq 1$ a.e.

# Section 2.8 – Sequences of Integrable Functions

## Problems

### 2.8.1

A measurable function $f$ is acalled a *null function* if $f = 0$ a.e. We shall say that $f$ is *equivalent* to $g$ (and write $f \sim g$) if $f - g$ is a null function. Denote by $\bar{f}$ the class of all measurable functions that are equivalent to $f$. We denote by $L^1(X, \mathcal{A}, \mu)$, or, more briefly, by $L^1(X, \mu)$, the set of all classes $\bar{f}$ for which $f$ is integrable, and define on it the function

$$\rho(\bar{f}, \bar{g}) = \rho(f, g) = \int |f - g| d\mu.$$

[Note that if $f_0 \in \bar{f}$, $g_0 \in \bar{g}$, then $\rho(f_0, g_0) = \rho(f, g)$.] Prove that $L^1(X, \mu)$ is a complete metric space with the metric $\rho$.

*Solution.* Let $(\bar{f}_n)$ be a Cauchy sequence in $L^1(X, \mu)$. Then

$$\int |f_m - f_n| d\mu = \rho(\bar{f}_m, \bar{f}_n) \to 0$$

as $m, n \to \infty$, so the sequence $(f_n)$ (of representative functions) is Cauchy in the mean. By Theorem 2.8.3 there is an integrable function $f$ such that $f_n \to f$ in the mean. Hence

$$\rho(\bar{f}_n, \bar{f}) = \int |f_n - f| d\mu \to 0$$

as $n \to \infty$. That is, $\bar{f}_n \to \bar{f}$ in $L^1(X, \mu)$.

### 2.8.2

TODO.

# Section 2.9 – Lebesgue's Bounded Convergence Theorem

## Problems

### 2.9.1

Let $f_n(x) = n$ if $0 \le x < 1/n$, $f_n(x) = 0$ if $1 \ge x \ge 1/n$. Then $\lim_n f_n(x) = 0$ on $[0, 1]$, but $\lim_n \int f_n d\mu = 1$ ($\mu$ is the Lebesgue integral). This example shows that the boundedness condition $|f_n| \le g$ in Theorem 2.9.1 is essential.

*Solution.* Indeed $f_n(x) \to 0$ for all $x > 0$, and

$$\mu([0,1] - (0,1]) = \mu(\{0\}) = 0,$$

so $f_n(x) \to 0$ a.e. Note that $f_n = n\chi_{[0,1/n)}$ for each $n$, so that

$$\int f_n d\mu = n\mu([0, 1/n)) = n\frac{1}{n} = 1.$$

Hence $\lim_n \int f_n d\mu = 1$.

### 2.9.4

If $\mu(X) < \infty$ and if $\{f_n\}$ is a sequence of measurable functions that converges uniformly to a function $f$, then $f$ is integrable and $\lim_n \int f_n d\mu = \int f d\mu$.

*Solution.* The problem statement, as written, is wrong. For a counterexample, let $X = \{1, 2, \dots\}$ and define the measure by $\mu(\{x\}) = 2^{-x}$ (extended additively to all subsets of $X$). Then $\mu(X) = 1$. Define $f$ by $f(x) = 2^x$ and let $f_n = f$ for all $n$. These definitions match the assumptions of the problem, but

$$f \ge g_n := \sum_{i=1}^{n} f(i)\chi_{\{i\}}$$

for all $n$, and

$$\int g_n d\mu = n \to \infty,$$

26

so $f$ cannot possibly be integrable.

We will therefore assume that the functions $f_n$ are integrable, not just measurable. Given any $\epsilon > 0$, there is an integer $N$ such that $n \geq N$ implies $|f_n - f| < \epsilon$. Hence

$$|f_m - f_n| \leq |f_m - f| + |f_n - f| < 2\epsilon$$

whenever $m, n \geq N$, and

$$\int |f_m - f_n| d\mu \leq 2\epsilon\mu(X).$$

It follows (since $\mu(X) < \infty$) that the sequence $(f_n)$ is Cauchy in the mean, hence that $f$ is integrable and $\int f d\mu = \lim_n \int f_n d\mu$ by Theorem 2.8.2

Remark: Assuming that the functions $f_n$ are integrable might not be the right way to fix to this problem, in particular since we did not even need to use the bounded convergence theorem to solve this version of it.

# Section 2.10 – Applications of Lebesgue's Bounded Convergence Theorem

## Problems

### 2.10.2

Derive the Lebesgue monotone convergence theorem from Fatou's Lemma.

*Solution.* Let $(f_n)$ be a monotone increasing sequence of non-negative integrable functions. Note that

$$\varliminf_n f_n(x) = \lim_n \inf_{j \geq n} f_j(x) = \lim_n f_n(x)$$

for all $x$, since $f_1(x) \leq f_2(x) \leq \cdots$. Similarly,

$$\varliminf_n \int f_n d\mu = \lim_n \int f_n d\mu.$$

By Fatou's lemma (Theorem 2.10.5),

$$\int \lim_n f_n d\mu = \int \varliminf_n f_n d\mu \leq \varliminf_n \int f_n d\mu = \lim_n \int f_n d\mu.$$

(Note that either the right-hand side or both sides of this inequality may be infinite.) On the other hand we also have

$$\lim_n \int f_n d\mu \leq \int \lim_n f_n d\mu$$

(again with the possibility of infinite values), since $f_j \leq \lim_n f_n$ for all $j$. Combining these results, we obtain

$$\lim_n \int f_n d\mu = \int \lim_n f_n d\mu,$$

with both sides either finite or infinite.

**2.10.7**

Let $\{f_n\}$ be a sequence of integrable functions. Prove that if $\sum_{n=1}^{\infty} \int |f_n| d\mu < \infty$, then the series $\sum_{n=1}^{\infty} f_n(x)$ is convergent to an integrable function $f(x)$, and

$$\int f \, d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

*Solution.* Assume that

$$\sum_{n=1}^{\infty} \int |f_n| \, d\mu < \infty.$$

For each $n \geq 1$, let

$$g_n = \sum_{i=1}^{n} f_i^{+}.$$

Then $(g_n)$ is a monotone-increasing sequence of nonnegative integrable functions. Also define $g = \lim_n g_n$. By the monotone convergence theorem (Theorem 2.10.4),

$$\int g \, d\mu = \lim_n \int g_n d\mu = \lim_n \int \sum_{i=1}^{n} f_i^{+} d\mu = \lim_n \sum_{i=1}^{n} \int f_i^{+} d\mu = \sum_{n=1}^{\infty} \int f_n^{+} d\mu,$$

which is finite, since

$$\sum_{n=1}^{\infty} \int f_n^{+} d\mu \leq \sum_{n=1}^{\infty} \int |f_n| \, d\mu < \infty.$$

Similarly defining $h_n = \sum_{i=1}^{n} f_i^{-}$ and $h = \lim_n h_n$, we find that

$$\int h \, d\mu = \sum_{n=1}^{\infty} \int f_i^{-} d\mu < \infty.$$

Next, define $f = g - h$. Note that $f$ is integrable by the above, and that

$$f = \sum_{n=1}^{\infty} (f_n^{+} - f_n^{-}) = \sum_{n=1}^{\infty} f_n.$$

Also by what we found above,

$$\int f \, d\mu = \int g \, d\mu - \int h \, d\mu = \sum_{n=1}^{\infty} \int f_n^{+} d\mu - \sum_{n=1}^{\infty} \int f_n^{-} d\mu,$$

from which we easily get

$$\int f \, d\mu = \sum_{n=1}^{\infty} \int f_n d\mu.$$

28

**2.10.8**

Let $\{f_n\}$ be a sequence of nonnegative integrable functions. Prove that if the series $f(x) = \sum f_n(x)$ is an integrable function, then $\sum_{n=1}^{\infty} \int f_n d\mu < \infty$.

*Solution.* Assume that $f$ is integrable. Then

$$\sum_{n=1}^{m} \int f_n d\mu = \int \sum_{n=1}^{m} f_n d\mu \leq \int f \, d\mu < \infty$$

for every $m$. Also,

$$\sum_{n=1}^{m} \int f_n d\mu \leq \sum_{n=1}^{m+1} \int f_n d\mu$$

for every $m$, since the functions $f_n$ are nonnegative. Hence the partial sums make up a monotone-increasing sequence of real numbers that is bounded above, and it follows that they converge to a finite number.

**2.10.9**

Let $f$ and $f_n$ $(n = 1, 2, \ldots)$ be integrable functions such that $0 \leq f_n(x) \leq f(x)$ a.e. Then

$$\int \left( \overline{\lim_{n}} f_n \right) d\mu \geq \overline{\lim_{n}} \int f_n d\mu \geq \underline{\lim_{n}} \int f_n d\mu \geq \int \left( \underline{\lim_{n}} f_n \right) d\mu.$$

*Solution.* For each $n \geq 1$, define

$$E_n = \{x; \, 0 \leq f_n(x) \leq f(x)\},$$

and let $N = \bigcup_n E_n^c$. Each set $E_n^c$ has measure zero by the given assumptions on $f_n$ and $f$, hence so does $N$. Define new functions $\tilde{f}_n$ and $\tilde{f}$ which are identical to $f_n$ and $f$ except that they vanish on $N$. Then $\tilde{f}_n(x) = f_n(x)$ a.e., $\tilde{f}(x) = f(x)$ a.e., and $0 \leq \tilde{f}_n(x) \leq \tilde{f}(x)$ *everywhere*. The new functions are integrable by Problem 2.6.1, and their integrals are identical to the old ones. Similarly, $\underline{\lim}_n \tilde{f}_n(x) = \underline{\lim}_n f_n(x)$ a.e. and $\overline{\lim}_n \tilde{f}_n(x) = \overline{\lim}_n f_n(x)$ a.e., once again with integrals unchanged (integrability follows from Theorem 2.10.1, with $f$ or $\tilde{f}$ bounding from above). Thus we can safely assume that $0 \leq f_n(x) \leq f(x)$ holds everywhere; otherwise we simply work with the functions $\tilde{f}_n$ and $\tilde{f}$ instead.

With such an assumption the inequality

$$\int \underline{\lim_{n}} f_n d\mu \leq \underline{\lim_{n}} \int f_n d\mu$$

follows immediately from Fatou's lemma. The next inequality,

$$\underline{\lim_{n}} \int f_n d\mu \leq \overline{\lim_{n}} \int f_n d\mu,$$

is trivial. For the final inequality, note that

$$f(x) - \overline{\lim_n} f_n(x) = \underline{\lim_n}(f(x) - f_n(x))$$

for all $x$. Hence, again by Fatou's lemma,

$$\int f\, d\mu - \int \overline{\lim_n} f_n d\mu = \int \underline{\lim_n}(f - f_n)d\mu$$

$$\leq \underline{\lim_n} \int (f - f_n)d\mu$$

$$= \underline{\lim_n} \left( \int f\, d\mu - \int f_n d\mu \right)$$

$$= \int f\, d\mu - \overline{\lim_n} \int f_n d\mu,$$

whereby

$$\overline{\lim_n} \int f_n d\mu \leq \int \overline{\lim_n} f_n d\mu.$$

**2.10.11**

Let $X = \bigcup_{n=1}^\infty E_n$, $E_n \subset E_{n+1}$ for all $n$. Let $f$ be a nonnegative measurable function. Prove that

$$\int f\, d\mu = \lim_{n\to\infty} \int_{E_n} f\, d\mu.$$

*Solution.* Note that $(\chi_{E_n} f)$ is a monotone-increasing sequence of nonnegative measurable functions, and that

$$\int_{E_n} f\, d\mu = \int \chi_{E_n} f\, d\mu.$$

Consider first the case that there is some $n$ such that $\chi_{E_n} f$ is not integrable. Then $f$ cannot be integrable, and $\chi_{E_m} f$ cannot be integrable for $m \geq n$, by Theorem 2.10.1 (since these functions bound $\chi_{E_n} f$ from above). Hence

$$\int f\, d\mu = \lim_n \int \chi_{E_n} f\, d\mu = \infty$$

in this case. If on the other hand all of the functions $\chi_{E_n} f$ are integrable, then the monotone convergence theorem tells us that

$$\lim_n \int \chi_{E_n} f\, d\mu = \int f\, d\mu,$$

since $\lim_n \chi_{E_n} f = f$.

**2.10.12**

Let $f$ be a nonnegative measurable function and let

$$f_n(x) = \begin{cases} f(x), & \text{if } f(x) \le n, \\ n, & \text{if } f(x) > n. \end{cases}$$

Prove that $\lim_n \int f_n d\mu = \int f d\mu$.

*Solution.* Note that $(f_n)$ is a monotone-increasing sequence of nonnegative measurable functions, with $\lim_n f_n = f$. If $f_n$ is not integrable for some $n$, then

$$\int f \, d\mu = \lim_n \int f_n d\mu = \infty$$

by Theorem 2.10.1. Otherwise the result follows by the monotone convergence theorem.

**2.10.14**

Give an example where Fatou's lemma holds with strict inequality.

*Solution.* Consider $[0, 1] \subset \mathbb{R}$ with Lebesgue measure. Let $f_n = \chi_{[0,1/2]}$ for odd $n$, and $f_n = \chi_{(1/2,1]}$ for even $n$. Then

$$\int f_n d\mu = \frac{1}{2}$$

for all $n$, so that

$$\varliminf_n \int f_n d\mu = \frac{1}{2},$$

but $\varliminf_n f_n = 0$ so that

$$\int \varliminf_n f_n d\mu = 0.$$

# Section 2.12 – The Radon-Nikodym Theorem

## Problems

### 2.12.1

Prove that if $f$ is integrable with respect to $\mu$ then it is also integrable with respect to $\mu^+$ and $\mu^-$. [Therefore the integrals in (2.12.5) are well defined.]

*Solution.* Since $f$ is $|\mu|$-integrable there exists a sequence $(f_n)$ of $|\mu|$-integrable simple functions that is Cauchy in the mean w.r.t. $|\mu|$ and converges to $f$ a.e. w.r.t. $|\mu|$. Note that $\mu^+ \le |\mu|$; consequently the functions $f_n$ are $\mu^+$-integrable. Another simple consequence is that $\int \varphi \, d\mu^+ \le \int \varphi \, d|\mu|$ for all nonnegative $|\mu|$-integrable simple functions $\varphi$. Hence

$$\int |f_m - f_n| \, d\mu^+ \le \int |f_m - f_n| \, d|\mu|$$

for all $m, n \geq 1$, and it follows that $(f_n)$ is Cauchy in the mean w.r.t. $\mu^+$. Finally, letting $N = \{x; \, f_n(x) \not\to f(x)\}$, we have $\mu^+(N) \leq |\mu|(N) = 0$, so that $f_n \to f$ a.e. w.r.t. $\mu^+$. Thus $f$ is $\mu^+$-integrable. The proof that $f$ is $\mu^-$-integrable is similar.

**2.12.2**

The Radon-Nikodym theorem remains true in case $\mu$ is a $\sigma$-finite signed measure.

*Solution.* Let $X = A \cup B$ be a Hahn decomposition of $(X, \mathcal{A}, \mu)$, and let $\mu = \mu^+ - \mu^-$ be the corresponding Jordan decomposition. Let $\mathcal{A}_A \subset \mathcal{A}$ be the collection of all measurable subsets of $A$, and $\mathcal{A}_B \subset \mathcal{A}$ the measurable subsets of $B$. We will consider separately the measure subspaces $(A, \mathcal{A}_A, \mu^+)$ and $(B, \mathcal{A}_B, \mu^-)$ (c.f. Problems 2.1.1-2.1.3).

Let $\nu$ be $\sigma$-finite measure on $\mathcal{A}$, with $\nu \ll \mu$. Note that $|\mu|(E) = \mu^+(E)$ for $E \in \mathcal{A}_A$, hence $\nu \ll \mu^+$ on $(A, \mathcal{A}_A, \mu^+)$. The Radon-Nikodym theorem yields a measurable function $f : A \to \mathbb{R}$ such that

$$\nu(E) = \int_E f \, d\mu^+ \quad \text{for all } E \in \mathcal{A}_A \text{ such that } |\nu|(E) < \infty.$$

Similarly, there is a measurable function $g : B \to \mathbb{R}$ such that

$$\nu(E) = \int_E g \, d\mu^- \quad \text{for all } E \in \mathcal{A}_B \text{ such that } |\nu|(E) < \infty.$$

It follows that

$$\nu(E) = \nu(E \cap A) + \nu(E \cap B) = \int_{E \cap A} f \, d\mu^+ + \int_{E \cap B} g \, d\mu^-$$

for all $E \in \mathcal{A}$ such that $|\nu|(E) < \infty$. Define $h : X \to \mathbb{R}$ such that $h = f$ on $A$ and $h = -g$ on $B$. Clearly $h$ is measurable, and if $E \in \mathcal{A}$ and $|\nu|(E) < \infty$, then $h$ is $\mu^+$-integrable on $E$ since it is separately integrable on $E \cap A$ (where it equals $f$) and on $E \cap B$ (since $\mu^+(E \cap B) = 0$, c.f. Theorem 2.7.4); similarly $h$ is $\mu^-$-integrable on $E$. Thus, for such $E$,

$$\begin{aligned}
\nu(E) &= \int_{E \cap A} h \, d\mu^+ - \int_{E \cap B} h \, d\mu^- \\
&= \int_{E \cap A} h \, d\mu^+ + \int_{E \cap B} h \, d\mu^+ - \int_{E \cap A} h \, d\mu^- - \int_{E \cap B} h \, d\mu^- \\
&= \int_E h \, d\mu^+ - \int_E h \, d\mu^- \\
&= \int_E h \, d\mu.
\end{aligned}$$

Finally, suppose there is another measurable function $\tilde{h} : X \to \overline{\mathbb{R}}$ (where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$) such that $\nu(E) = \int_E \tilde{h} \, d\mu$ whenever $E \in \mathcal{A}$ and $|\nu|(E) < \infty$.

Define $\tilde{f} : A \to \overline{\mathbb{R}}$ and $\tilde{g} : B \to \overline{\mathbb{R}}$ such that they equal $\tilde{h}$ on their respective domains. Then

$$\nu(E) = \int_E \tilde{h} \, d\mu = \int_E \tilde{h} \, d\mu^+ + \int_E \tilde{h} \, d\mu^- = \int_E \tilde{f} \, d\mu^+$$

for all $E \in \mathcal{A}_A$ such that $|\nu|(E) < \infty$, hence $\tilde{f} = f$ a.e. in $(A, \mathcal{A}_A, \mu^+)$ by the Radon-Nikodym theorem. Similarly, $\tilde{g} = g$ a.e. in $(B, \mathcal{A}_B, \mu^-)$. Let $N = \{x; \tilde{h}(x) \neq h(x)\}$, and note that $N \cap A = \{x \in A; \tilde{f}(x) \neq f(x)\}$ and $N \cap B = \{x \in B; \tilde{g}(x) \neq g(x)\}$. Then

$$|\mu|(N) = \mu^+(N) + \mu^-(N) = \mu^+(N \cap A) + \mu^-(N \cap B) = 0,$$

so $\tilde{h} = h$ a.e. w.r.t. $|\mu|$.

**2.12.3**

If $\nu$ and $\mu$ are $\sigma$-finite signed measures and $\nu \ll \mu$, then the set $\{x; (d\nu/d\mu)(x) = 0\}$ has $\nu$-measure zero.

*Solution.* By the preceding problem, there exists a Radon-Nikodym derivative $f$ of $\nu$ w.r.t. $\mu$. Let $N = \{x; f(x) = 0\}$. Since $\nu$ is $\sigma$-finite we can write $X = \bigcup_n A_n$ with $A_n$ disjoint measurable sets such that $|\nu|(A_n) < \infty$. It follows that

$$\nu(N) = \sum_n \nu(N \cap A_n) = \sum_n \int_{N \cap A_n} f \, d\mu = \sum_n 0 = 0.$$

**2.12.4**

Let $\nu$ and $\mu$ be $\sigma$-finite signed measures and let $\nu \ll \mu$. For any $\nu$-integrable function $\varphi$, the function $\varphi(d\nu/d\mu)$ is $\mu$-integrable and

$$\int \varphi \, d\nu = \int \varphi \frac{d\nu}{d\mu} \, d\mu.$$

[*Hint:* Consider first the case where $\mu$ and $\nu$ are measures and $\varphi$ is a simple function.]

*Solution.* We will divide the proof into multiple parts.

(i) Consider first the case where $\mu$ and $\nu$ are measures (i.e. nonnegative) and $\varphi$ is an integrable simple function. Then

$$\varphi = \sum_{i=1}^n \alpha_i \chi_{E_i}$$

for some measurable sets $E_i$ with $\nu(E_i) < \infty$. It follows by the Radon-Nikodym theorem that

$$\int \varphi \, d\nu = \sum_{i=1}^n \alpha_i \nu(E_i) = \sum_{i=1}^n \alpha_i \int_{E_i} \frac{d\nu}{d\mu} d\mu = \int \sum_{i=1}^n \alpha_i \chi_{E_i} \frac{d\nu}{d\mu} d\mu$$
$$= \int \varphi \frac{d\nu}{d\mu} d\mu.$$

In particular, $\varphi \cdot d\nu/d\mu$ is $\mu$-integrable.

(ii) Next, let $\mu$ and $\nu$ remain measures, but let $\varphi$ be any nonnegative $\nu$-integrable function. Theorem 2.2.5 provides a monotone-increasing sequence $(\varphi_n)$ of nonnegative simple functions such that $\varphi_n(x) \to \varphi(x)$ everywhere, and each $\varphi_n$ is $\nu$-integrable since $\varphi_n \leq \varphi$ (Theorem 2.10.1). By the monotone convergence theorem and part (i) above,

$$\int \varphi \, d\nu = \lim_n \int \varphi_n \, d\nu = \lim_n \int \varphi_n \frac{d\nu}{d\mu} d\mu.$$

Let $N = \{x; (d\nu/d\mu)(x) < 0\}$. By $\sigma$-finiteness there are measurable sets $A_n$ such that $X = \bigcup_n A_n$ and $\nu(A_n) < \infty$. For each $n \in \mathbb{N}$ we have

$$0 \leq \nu(N \cap A_n) = \int_{N \cap A_n} \frac{d\nu}{d\mu} d\mu \leq 0,$$

hence $\mu(N \cap A_n) = 0$ by Theorem 2.7.5 (applied to $-d\nu/d\mu$). It follows that $\mu(N) = 0$, i.e. that $d\nu/d\mu \geq 0$ $\mu$-a.e. We can therefore redefine $d\nu/d\mu$ to be 0 on $N$, and it will still be a Radon-Nikodym derivative.

The non-negative integrable functions $\varphi_n \cdot d\nu/d\mu$ make up a monotone-increasing sequence with pointwise limit $\varphi \cdot d\nu/d\mu$. Thus

$$\lim_n \int \varphi_n \frac{d\nu}{d\mu} d\mu = \int \varphi \frac{d\nu}{d\mu} d\mu$$

by monotone convergence, yielding

$$\int \varphi \, d\nu = \int \varphi \frac{d\nu}{d\mu} d\mu.$$

(iii) Let $\mu$ be a measure, $\nu$ a signed measure, and $\varphi$ a nonnegative $\nu$-integrable function. Then

$$\int \varphi \, d\nu = \int \varphi \, d\nu^+ - \int \varphi \, d\nu^- = \int \varphi \frac{d\nu^+}{d\mu} d\mu - \int \varphi \frac{d\nu^-}{d\mu} d\mu$$

by part (ii). Note that

$$\nu(E) = \nu^+(E) - \nu^-(E) = \int \frac{d\nu^+}{d\mu} d\mu - \int \frac{d\nu^-}{d\mu} d\mu = \int \left( \frac{d\nu^+}{d\mu} - \frac{d\nu^-}{d\mu} \right) d\mu$$

for all measurable subsets $E$ with $|\nu|(E) < \infty$. Hence

$$\left( \frac{d\nu^+}{d\mu} - \frac{d\nu^-}{d\mu} \right) = \frac{d\nu}{d\mu} \quad \mu\text{-a.e.}$$

by the Radon-Nikodym theorem, and it follows that

$$\int \varphi \, d\nu = \int \varphi \frac{d\nu}{d\mu} d\mu.$$

(iv) Now let both $\mu$ and $\nu$ be signed measures, with $\varphi$ still a nonnegative $\nu$-integrable function. As in the solution of Problem 2.12.2, let $X = A \cup B$ be a Hahn decomposition of $(X, \mathcal{A}, \mu)$. We obtain Radon-Nikodym derivatives $d\nu/d\mu^+$ and $d\nu/d\mu^-$ on the measure subspaces $(A, \mathcal{A}_A, \mu^+)$ and $(A, \mathcal{A}_B, \mu^-)$, respectively, and extend them to $X$ by setting $d\nu/d\mu^+ = 0$ on $B$ and $d\nu/d\mu^- = 0$ on $A$. Then

$$
\begin{aligned}
\int \varphi \, d\nu &= \int_A \varphi \, d\nu + \int_B \varphi \, d\nu \\
&= \int \varphi \frac{d\nu}{d\mu^+} \, d\mu^+ + \int \varphi \frac{d\nu}{d\mu^+} \, d\mu^+ \\
&= \int \varphi \left( \frac{d\nu}{d\mu^+} - \frac{d\nu}{d\mu^-} \right) d\mu^+ - \int \varphi \left( \frac{d\nu}{d\mu^+} - \frac{d\nu}{d\mu^-} \right) d\mu^- \\
&= \int \varphi \left( \frac{d\nu}{d\mu^+} - \frac{d\nu}{d\mu^-} \right) d\mu,
\end{aligned}
$$

with the second equality following from part (iii). In particular,

$$
\nu(E) = \int \chi_E \, d\nu = \int \chi_E \left( \frac{d\nu}{d\mu^+} - \frac{d\nu}{d\mu^-} \right) d\mu
$$

for all $E \in \mathcal{A}$ with $|\nu|(E) < \infty$, and it follows by Problem 2.12.2 that

$$
\frac{d\nu}{d\mu} = \frac{d\nu}{d\mu^+} - \frac{d\nu}{d\mu^-} \quad \mu\text{-a.e.}
$$

Thus

$$
\int \varphi \, d\nu = \int \varphi \frac{d\nu}{d\mu} \, d\mu.
$$

(v) Finally, let $\mu$ and $\nu$ be signed, and $\varphi$ be $\nu$-integrable. The result follows by applying the conclusion of part (iv) to $\varphi^+$ and $\varphi^-$.

# Chapter 3 – Metric Spaces

## Section 3.1 – Topological and Metric Spaces

### Problems

**3.1.1**

Prove that if $(X, p)$ is a metric space, and if

$$\hat{\rho}(x, y) = \frac{\rho(x, y)}{1 + \rho(x, y)},$$

then also $(X, \hat{\rho})$ is a metric space. [*Hint:* Cf. the proof of (3.1.3).]

*Solution.* The only nonobvious property is the triangle inequality. Let $x, y, z$ be arbitrary points of $X$. Since $t \mapsto t/(1+t)$ is monotone increasing on $[0, \infty)$, and since $\rho(x, z) \leq \rho(x, y) + \rho(y, z)$, we have

$$\frac{\rho(x, z)}{1 + \rho(x, z)} \leq \frac{\rho(x, y) + \rho(y, z)}{1 + \rho(x, y) + \rho(y, z)}.$$

Moreover, by equation (3.1.3),

$$\frac{\rho(x, y) + \rho(y, z)}{1 + \rho(x, y) + \rho(y, z)} \leq \frac{\rho(x, y)}{1 + \rho(x, y)} + \frac{\rho(y, z)}{1 + \rho(y, z)}.$$

It follows that $\hat{\rho}(x, z) \leq \hat{\rho}(x, y) + \hat{\rho}(y, z)$.

**3.1.2**

Let $X, \rho, \hat{\rho}$ be as in Problem 3.1.1. Prove that $\rho(x_n, x) \to 0$ if and only if $\hat{\rho}(x_n, x) \to 0$. Give an example showing that $\rho$ and $\hat{\rho}$ are not equivalent in general.

*Solution.* It is clear that $\rho(x_n, x) \to 0$ implies $\hat{\rho}(x_n, x) \to 0$, since $\hat{\rho} \leq \rho$.

Conversely, suppose $\hat{\rho}(x_n, x) \to 0$. Given any $\epsilon > 0$, there is a positive integer $N$ such that

$$\hat{\rho}(x_n, x) < \frac{\epsilon}{1 + \epsilon} \quad (n \geq N).$$

Substituting the definition of $\hat{\rho}$ and rearranging yields $\rho(x_n, x) < \epsilon$.

If $\rho$ and $\hat{\rho}$ are equivalent then, in particular, there exists a positive constant $\beta$ such that

$$\frac{\rho(x, y)}{\hat{\rho}(x, y)} \leq \beta$$

whenever $x \neq y$. But

$$\frac{\rho(x, y)}{\hat{\rho}(x, y)} = 1 + \rho(x, y),$$

so this is impossible if $X$ is unbounded (w.r.t. $\rho$), say if $X = \mathbb{R}^n$ and $\rho$ is the Euclidean metric.

### 3.1.6

Prove that the spaces $l^1, s, c, c_0$ are separable metric spaces.

*Solution.* For each of the spaces $X$ we will take an arbitrary element $x \in X$ and demonstrate that every $\epsilon$-ball around $x$ contains a point $y$ of a certain countable subset $Y \subset X$. It will then follow that $Y$ is dense in $X$, and hence that $X$ is separable.

Let $x = (x_i) \in l^1$. Fix $\epsilon > 0$. Since $\sum_i |x_i| < \infty$, there exists $n$ such that

$$\sum_{i=n+1}^{\infty} |x_i| < \frac{\epsilon}{2}.$$

For $i = 1, \ldots, n$, choose $y_i \in \mathbb{Q}$ (or $y_i \in \mathbb{Q} + i\mathbb{Q}$ in the complex case) such that $|x_i - y_i| < \epsilon/2n$, and let $y = (y_1, \ldots, y_n, 0, 0, \ldots)$. Then

$$\rho(x, y) = \sum_{i=1}^{n} |x_i - y_i| + \sum_{i=n+1}^{\infty} |x_i| < \epsilon.$$

Moreover, $y$ is an element of the subset of $l^1$ consisting of sequences with rational components, with only finitely many being nonzero. This subset is easily seen to be countable, and it follows that $l^1$ is separable.

Next, let $x = (x_i) \in s$, and fix $\epsilon > 0$. Choose $n$ such that

$$\sum_{i=n+1}^{\infty} \frac{1}{2^i} \frac{|x_i|}{1 + |x_i|} < \frac{\epsilon}{2}.$$

For $i = 1, \ldots, n$, choose $y_i \in \mathbb{Q}$ such that $|x_i - y_i| < \epsilon/2$, and let $y = (y_1, \ldots, y_n, 0, 0, \ldots)$. Then

$$\rho(x, y) = \sum_{i=1}^{n} \frac{1}{2^i} \frac{|x_i - y_i|}{1 + |x_i - y_i|} + \sum_{i=n+1}^{\infty} \frac{1}{2^i} \frac{|x_i|}{1 + |x_i|} < \epsilon.$$

Similarly to the above, it follows that $s$ is dense.

Finally, let $x = (x_i) \in c$. (This argument will also cover $c_0$.) Let $\xi = \lim_i x_i$ and fix $\epsilon > 0$. Choose $n$ such that $|x_i - \xi| < \epsilon/2$ for all $i \geq n$. For $i = 1, \ldots, n-1$,

37

choose $y_i \in \mathbb{Q}$ such that $|x_i - y_i| < \epsilon$. Also choose $\eta \in \mathbb{Q}$ such that $|\xi - \eta| < \epsilon/2$ (take $\eta = \xi = 0$ in the $c_0$ case), so that

$$|x_i - \eta| \leq |x_i - \xi| + |\xi - \eta| < \epsilon$$

for all $i \geq n$. Let $y = (y_1, \ldots, y_{n-1}, \eta, \eta, \ldots)$. Then

$$\rho(x, y) = \sup_i |x_i - y_i| \leq \epsilon.$$

It follows that $c$ is separable (and $c_0$ as well).

### 3.1.10

If $\rho(x_n, x) \to 0$, $\rho(y_n, y) \to 0$, then $\rho(x_n, y_n) \to \rho(x, y)$.

*Solution.* The triangle inequality yields

$$\rho(x_n, y_n) \leq \rho(x_n, x) + \rho(x, y) + \rho(y, y_n)$$

and

$$\rho(x, y) \leq \rho(x, x_n) + \rho(x_n, y_n) + \rho(y_n, y).$$

Hence

$$|\rho(x_n, y_n) - \rho(x, y)| \leq \rho(x_n, x) + \rho(y_n, y),$$

and the result follows.

## Section 3.2 − $L^p$ Spaces

### Problems

#### 3.2.4

Prove that $l^p$ is separable if $1 \leq p < \infty$.

*Solution.* The argument is of the same sort as in Problem 3.1.6. Given $x = (x_i) \in l^p$ ($1 \leq p < \infty$) and $\epsilon > 0$, choose $n$ such that

$$\sum_{i=n+1}^{\infty} |x_i|^p < \frac{\epsilon^p}{2}.$$

(This is possible since $\sum_i |x_i|^p < \infty$ for $x \in l^p$.) For $i = 1, \ldots, n$, choose $y_i \in \mathbb{Q}$ such that $|x_i - y_i|^p < \epsilon^p/2n$ and let $y = (y_1, \ldots, y_n, 0, 0, \ldots)$. Then $\|x - y\|_p < \epsilon$, and the conclusion follows as in Problem 3.1.6.

**3.2.5**

Prove that $l^p$ is not a metric space if $0 < p < 1$.

*Solution.* Let $x = (1, 0, 0, \dots)$, $y = (0, 0, \dots)$, $z = (0, 1, 0, 0, \dots)$. Then

$$\|x - z\|_p = 2^{1/p}$$

while

$$\|x - y\|_p + \|y - z\|_p = 2.$$

But $2^{1/p} > 2$ if $0 < p < 1$, so the triangle inequality does not hold.

**3.2.6**

Prove that the space $C[a, b]$ with the metric

$$\rho(f, g) = \int_a^b |f(t) - g(t)|\, dt$$

is not a complete metric space.

*Solution.* Note that the integral defining $\rho$ is guaranteed to exist by Theorem 2.11.1 and the extreme value theorem from elementary analysis, and it can be taken in the Riemann sense. It is easily seen that $\rho$ is indeed a metric.

For simplicity, let us assume $a = 0$ and $b = 1$. (The general argument is identical modulo a coordinate transformation.) Define a sequence of functions in $C[0, 1]$ by $f_n(x) = x^n$. Clearly $(f_n)$ converges pointwise to the discontinuous function

$$f(x) = \begin{cases} 0 & (0 \le x < 1), \\ 1 & (x = 1). \end{cases}$$

For $n \le m$ we have $f_n \ge g_n$, hence

$$\rho(f_n, f_m) = \int_0^1 f_n(t)\, dt - \int_0^1 f_m(t)\, dt = \frac{1}{n+1} - \frac{1}{m+1},$$

which goes to zero as $n, m \to \infty$. Thus $(f_n)$ is a Cauchy sequence in $C[0, 1]$ which converges to a function $f$ which is not in $C[0, 1]$, so $(C[0, 1], \rho)$ is not a complete metric space.

# Section 3.4 – Complete Metric Spaces

## Problems

**3.4.5**

A set $Y$ in a metric space $X$ is said to be *of the first category in $X$* if it is contained in a countable union of nowhere dense sets of $X$. If $Y$ is not of the first category in $X$, then it is said to be *of the second category in $X$*. The real line with the Euclidean metric is a space of the second category. Prove, however, that, as a subset of the Euclidean plane, the real line is a set of the first category.

*Solution.* We identify the real line with the subset $L = \{(x, y) \in \mathbb{R}^2 : y = 0\}$, with the induced metric topology. Indeed $L$ is homeomorphic to $\mathbb{R}$ via the map $(x, 0) \mapsto x$, as can easily be shown.

The sets of the form $B((0,0), r) \cap L = \{(x, 0) : |x| < r\}$ are easily seen to be nowhere dense in $\mathbb{R}^2$, and we have

$$L = \bigcup_{n=1}^{\infty} B((0,0), n) \cap L.$$

Hence $L$ is of the first category in $\mathbb{R}^2$.

### 3.4.7

Let $f(x)$ be a real-valued function on the real line. Prove that there is a nonempty interval $(a, b)$ and a positive number $c$ such that for any $x \in (a, b)$ there is a sequence $\{x_n\}$ such that $x_n \to x$ and $|f(x_n)| \leq c$.

*Solution.* Note that

$$\mathbb{R} = |f|^{-1}(\mathbb{R}) = |f|^{-1}\left(\bigcup_{n=1}^{\infty} (-\infty, n]\right) = \bigcup_{n=1}^{\infty} |f|^{-1}\left((-\infty, n]\right).$$

Since $\mathbb{R}$ is of the second category (by Theorem 3.4.2), the sets $|f|^{-1}((-\infty, n])$ cannot all be nowhere dense. Hence there is a positive integer $c$ such that the closure of $|f|^{-1}((-\infty, c])$ has nonempty interior. On account of being a nonempty open set, this interior contains an open interval $(a, b)$.

Let $x \in (a, b)$. Then, since $x$ is contained in the closure of $|f|^{-1}((-\infty, c])$, every neighborhood of $x$ contains a point of $|f|^{-1}((-\infty, c])$. It follows that we can construct a sequence $(x_n)$ in $|f|^{-1}((-\infty, c])$ converging to $x$, and this sequence will satisfy $|f(x_n)| \leq c$ for all $n$.

## Section 3.5 – Compact Metric Spaces

### Problems

### 3.5.4

A subset $F$ of a compact metric space is compact if and only if it is closed.

*Solution.* Let $X$ be a compact metric space, and $F$ a subset of $X$. If $F$ is compact then it is also closed by Corollary 3.5.5.

Conversely, suppose that $F$ is closed. Let $\mathscr{C}$ be any open cover of $F$. The collection $\mathscr{C} \cup \{X - F\}$ is an open cover of $X$, hence it has a finite subcover, consisting of some sets $E_1, \ldots, E_n \in \mathscr{C}$ and perhaps $\{X - F\}$. The sets $E_1, \ldots, E_n$ cover $F$, hence $F$ is compact.

**3.5.5**

A subset $Y$ of a metric space is totally bounded if and only if its closure $\overline{Y}$ is totally bounded.

*Solution.* Clearly $Y$ is totally bounded whenever $\overline{Y}$ is. Conversely, suppose $Y$ is totally bounded, and let $\epsilon > 0$ be given. Then $Y$ admits a finite $\epsilon/2$-covering,

$$\{B(x_1, \epsilon/2), \ldots, B(x_n, \epsilon/2)\}.$$

The union of the closed balls,

$$\bigcup_{i=1}^{n} \overline{B}(x_i, \epsilon/2),$$

is a closed set containing $Y$, hence it contains $\overline{Y}$. It follows that

$$\overline{Y} \subset \bigcup_{i=1}^{n} \overline{B}(x_i, \epsilon/2) \subset \bigcup_{i=1}^{n} B(x_i, \epsilon),$$

so $\overline{Y}$ admits a finite $\epsilon$-covering.

**3.5.6**

The intersection of any number of compact subsets of a metric space is a compact space.

*Solution.* Let $\mathscr{C}$ be a collection of compact subsets of a metric space $X$. By Corollary 3.5.5 each $C \in \mathscr{C}$ is closed, hence the intersection $K = \bigcap \mathscr{C}$ is closed as well. Let $C'$ be some particular member of $\mathscr{C}$. Then $K$ is a closed subset of the compact space $C'$, thus itself compact by Problem 3.5.4.

Remark: Really the above shows that $K$ is compact *in* $C'$. However, one easily proves the general result that if $F$ is compact in $G$, and $G$ is compact in $H$, then $F$ is compact in $H$. Hence $K$ is indeed compact in $X$.

**3.5.7**

A subset of a separable metric space is a separable space.

*Solution.* Let $X$ be a separable metric space, and $Y \subset X$ an arbitrary subset. Since $X$ is separable, there exists a countable subset $A \subset X$ such that $\overline{A} = X$.

For each $n \in \mathbb{N}$, define a countable subset $Z_n \subset Y$ as follows. For each $x \in A$ such that $B(x, 1/(2n)) \cap Y \neq \emptyset$, choose an element $z \in B(x, 1/(2n)) \cap Y$, and let $Z_n$ be the set of these elements. Then the collection $\{B(z, 1/n) : z \in Z_n\}$ covers $Y$. Indeed, if $y \in Y$, then there is some $x \in A$ such that $d(y, x) < 1/(2n)$, hence $B(x, 1/(2n)) \cap Y \neq \emptyset$ and there is $z \in Z_n$ with $d(x, z) < 1/(2n)$, so that finally

$$d(y, z) \leq d(y, x) + d(x, z) < \frac{1}{n}.$$

The union $Z = \bigcup_n Z_n$ is a countable subset of $Y$, and by the above each point of $Y$ is within $1/n$ of a point of $Z$ for every $n \in \mathbb{N}$, thus $Y \subset \overline{Z}$ so that $Y$ is separable.

**3.5.8**

Show that a metric space is compact if and only if it has the following property: for every collection of closed subsets $\{F_\alpha\}$, if any finite subcollection has nonempty intersection, then the whole collection has a nonempty intersection.

*Solution.* The wording of the problem is potentially misleading. Replace "any" with "every" to make it nonambiguous. Also, let us introduce some useful terminology. A collection of subsets of a topological space is said to have the *finite intersection property* iff every finite subcollection has nonempty intersection. Our task is therefore to show that a metric space $X$ is compact iff every collection $\mathscr{F}$ of closed subsets of $X$ having the finite intersection property has nonempty intersection. However, the proof we will use is valid for general topological spaces, not just metric spaces.

Let $X$ be a topological space, and suppose first that $X$ is compact. Let $\mathscr{F}$ be a collection of closed subsets of $X$ with the finite intersection property. Assume towards a contradiction that $\bigcap \mathscr{F} = \emptyset$. Then

$$\bigcup_{F \in \mathscr{F}} (X - F) = X - \bigcap_{F \in \mathscr{F}} F = X,$$

so $\{X - F; \, F \in \mathscr{F}\}$ is an open cover of $X$. Since $X$ is compact, there are sets $F_1, \ldots, F_n \in \mathscr{F}$ such that

$$X = \bigcup_{i=1}^{n} (X - F_i) = X - \bigcap_{i=1}^{n} F_i.$$

But then $\bigcap_{i=1}^{n} F_i = \emptyset$, contradicting the hypothesis that $\mathscr{F}$ has the finite intersection property. We conclude that $\bigcap \mathscr{F} \neq \emptyset$.

Conversely, suppose that $X$ has the property described in the problem statement: every collection of closed subsets of $X$ with the finite intersection property has nonempty intersection. Let $\mathscr{U}$ be any open cover of $X$, and assume towards a contradiction that $\mathscr{U}$ has no finite subcover. Then

$$\bigcap_{i=1}^{n} (X - U_i) = X - \bigcup_{i=1}^{n} U_i \neq \emptyset$$

for every finite subcollection $\{U_1, \ldots, U_n\} \subset \mathscr{U}$. Hence $\mathscr{F} = \{X - U; \, U \in \mathscr{U}\}$ is a collection of closed subsets with the finite intersection property. By hypothesis $\mathscr{F}$ has nonempty intersection, so

$$\emptyset = X - \bigcup_{U \in \mathscr{U}} U = \bigcap_{U \in \mathscr{U}} (X - U) = \bigcap_{F \in \mathscr{F}} F \neq \emptyset,$$

a contradiction. We conclude that $\mathscr{U}$ does indeed have a finite subcover, and that $X$ is compact.

# Chapter 4 − Elements of Functional Analysis in Banach Spaces

## Section 4.1 − Linear Normed Spaces

### Problems

**4.1.4**

If $\{x_n\}$ is a convergent sequence in a normed linear space, with limit $x$, then also the sequence with elements $(x_1 + \cdots + x_n)/n$ is convergent to $x$.

*Solution.* For $n = 1, 2, \ldots$, define

$$\sigma_n = \frac{x_1 + \cdots + x_n}{n} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Let $\epsilon > 0$ be given. Since $(x_n)$ is convergent there exists a positive integer $m$ such that $\|x_n - x\| < \epsilon/2$ for all $n \geq m + 1$. For such $n$ we have

$$
\begin{aligned}
\|\sigma_n - x\| &= \left\| \frac{1}{n} \sum_{i=1}^{n} (x_i - x) \right\| \\
&\leq \frac{1}{n} \left\| \sum_{i=1}^{m} (x_i - x) \right\| + \frac{1}{n} \left\| \sum_{i=m+1}^{n} (x_i - x) \right\| \\
&\leq \frac{1}{n} \left\| \sum_{i=1}^{m} (x_i - x) \right\| + \frac{1}{n} \sum_{i=m+1}^{n} \|x_i - x\| \\
&\leq \frac{1}{n} \left\| \sum_{i=1}^{m} (x_i - x) \right\| + \frac{n - m - 1}{n} \cdot \frac{\epsilon}{2}.
\end{aligned}
$$

The first term on the right-hand side goes to zero as $n \to \infty$, and the other goes to $\epsilon/2$. Hence $\|\sigma_n - x\| < \epsilon$ for sufficiently large $n$, and it follows that $\sigma_n \to x$.

**4.1.6**

A normed linear space is a Banach space if the following property is satisfied: every absolutely convergent series is convergent.

*Solution.* Let $X$ be a normed linear space with the property that every absolutely convergent series is convergent. Let $(x_n)$ be a Cauchy sequence in $X$, and choose integers $N_1 < N_2 < \cdots$ such that $m, n \geq N_k$ implies $\|x_m - x_n\| < 2^{-k}$. Define a new sequence $(y_k)$ by $y_1 = x_{N_1}$ and $y_k = x_{N_k} - x_{N_{k-1}}$ for $k > 1$. Then

$$\sum_{k=1}^{\infty} \|y_k\| = \|x_{N_1}\| + \sum_{k=2}^{\infty} \|x_{N_k} - x_{N_{k-1}}\| \leq \|x_{N_1}\| + \sum_{k=1}^{\infty} 2^{-k} = \|x_{N_1}\| + 1.$$

It follows by our hypothesis on $X$ that

$$x_{N_k} = \sum_{i=1}^{k} y_i$$

converges to some $x \in X$. Given $\epsilon > 0$, choose an integer $j \geq 1$ such that $2^{-j} < \epsilon$, and $k > j$ such that

$$\|x_{N_k} - x\| < \epsilon - 2^{-j}.$$

Then, for all $n \geq N_k$,

$$\|x_n - x\| \leq \|x_n - x_{N_k}\| + \|x_{N_k} - x\| < 2^{-k} + \epsilon - 2^{-j} < \epsilon.$$

We conclude that every Cauchy sequence in $X$ is convergent, and hence that $X$ is a Banach space.

# Section 4.2 – Subspaces and Bases

## Problems

### 4.2.6

If a linear vector space is infinite-dimensional, then there exist on it norms that are not equivalent. [*Hint:* Let $\{y_\alpha\}$ be a Hamel basis and define norms by $\|x\|^2 = \sum_\alpha c_\alpha |\lambda_\alpha|^2$, where $x$ has the form (4.2.2) and $c_\alpha$ are positive numbers.]

*Solution.* Let $\{y_\alpha\}_{\alpha \in I}$ be a Hamel basis for an infinite-dimensional linear space $X$. For every $x = \sum_{\alpha \in I} \lambda_\alpha y_\alpha$, define

$$\|x\|_1 = \left( \sum_{\alpha \in I} |\lambda_\alpha|^2 \right)^{1/2}.$$

Then $\| \ \|_1$ is easily seen to be a norm on $X$.

Let $\{\alpha_i\}_{i=1}^{\infty}$ be a countable subset of the index set $I$. Define a set $\{c_\alpha\}_{\alpha \in I}$ of positive constants by $c_{\alpha_i} = 2^{-i}$ for $i = 1, 2, \ldots,$ and $c_\alpha = 1$ for $\alpha \notin \{\alpha_i\}$. Define another norm $\| \ \|_2$ by

$$\|x\|_2 = \left( \sum_{\alpha \in I} c_\alpha |\lambda_\alpha|^2 \right)^{1/2},$$

for $x = \sum_{\alpha \in I} \lambda_\alpha y_\alpha$.

To see that these two norms are not equivalent, note that

$$\|y_{\alpha_i}\|_2^2 = c_{\alpha_i} = 2^{-i} = 2^{-i} \|y_{\alpha_i}\|_1^2$$

for all $i$. Hence there exists no $\beta > 0$ such that $\|x\|_1 \leq \beta \|x\|_2$ for all $x \in X$.

# Section 4.3 – Finite-Dimensional Normed Linear Spaces

## Problems

### 4.3.1

Let $X$ be a finite-dimensional linear space. Then any two norms on $X$ are equivalent. (According to Problem 4.2.6, the assertion is false if $X$ is infinite-dimensional.)

*Solution.* Let $e_1, \ldots, e_n$ be a basis of $X$. For every $x = \sum_{i=1}^{n} \lambda_i e_i \in X$, let

$$\|x\|_1 = \sum_{i=1}^{n} |\lambda_i|.$$

It is easily verified that $\| \ \|_1$ is a norm on $X$. We will show that every norm on $X$ is equivalent to $\| \ \|_1$, and hence (by transitivity) that any two norms on $X$ are equivalent.

Given an arbitrary norm $\| \ \|$, we must prove the existence of positive constants $\alpha$ and $\beta$ such that

$$\alpha \|x\|_1 \leq \|x\| \leq \beta \|x\|_1$$

for all $x \in X$. These inequalities hold trivially for $x = 0$, so it suffices to consider nonzero $x$. In fact it is sufficient to consider $x$ in the "$\| \ \|_1$-sphere" $S_1 = \{x \in X : \|x\|_1 = 1\}$, where the inequalities reduce to

$$\alpha \leq \|x\| \leq \beta.$$

The inequality for general, nonzero $x$ then follows upon division by $\|x\|_1$.

We start by showing that that the map $x \mapsto \|x\|$ is continuous with respect to the metric $\rho_1(x, y) = \|x - y\|_1$. Let $\epsilon > 0$ be given, and write $M = \max(\|e_1\|, \ldots, \|e_n\|)$. Given

$$x = \sum_{i=1}^n \lambda_i e_i, \quad y = \sum_{i=1}^n \mu_i e_i,$$

satisfying $\rho_1(x, y) < \epsilon/M$, we have

$$\|x - y\| \le \sum_{i=1}^n |\lambda_i - \mu_i| \, \|e_i\| \le \sum_{i=1}^n |\lambda_i - \mu_i| \, M = \rho_1(x, y) M < \epsilon,$$

and continuity follows.

The sphere $S_1$ is obviously closed and bounded under $\|\ \|_1$, hence compact by Theorem 4.3.3. By Theorem 3.6.2 and the continuity established above, the map $x \mapsto \|x\|$ attains a maximum and a minimum on $S_1$. Let

$$\alpha = \inf_{x \in S_1} \|x\|, \quad \beta = \sup_{x \in S_1} \|x\|,$$

and note that $\alpha > 0$ since $\|x\| = \alpha$ is attained for some nonzero $x$. It follows that $\alpha$ and $\beta$ are positive constants such that $\alpha \le \|x\| \le \beta$ for all $x \in S_1$, so we are done.

### 4.3.2

Let $Y$ be a finite-dimensional linear subspace of a normed linear space $X$, and let $x_0 \in X$, $x_0 \notin Y$. Then there exists a point $y_0 \in Y$ such that

$$\inf_{y \in Y} \|x_0 - y\| = \|x_0 - y_0\|.$$

*Solution.* Note that $Y$ is closed by Theorem 4.3.2. Let $L = \inf_{y \in Y} \|x_0 - y\|$. For $n = 1, 2, \ldots$, choose $y_n \in Y$ such that

$$\|x_0 - y_n\| < L + \frac{1}{n}.$$

Note that $y_n \in B(x_0, L + 1) \cap Y$ for all $n$. This is a bounded subset of $Y$, so by Theorems 4.3.3 and 3.5.4, the sequence $(y_n)$ has a subsequence $(y_{n_k})$ converging to a point

$$y_0 \in \overline{B(x_0, L + 1) \cap Y} \subset \overline{Y} = Y.$$

For all $k$ we have

$$\|x_0 - y_0\| \le \|x_0 - y_{n_k}\| + \|y_{n_k} - y_0\|.$$

The left-hand side of this inequality is bounded below by $L$, and the right-hand side converges to $L$ as $k \to \infty$. It follows that $\|x_0 - y_0\| = L$.

**4.3.3**

A norm $\| \ \|$ is said to be *strictly convex* if $\|x\| = 1$, $\|y\| = 1$, $\|x + y\| = 2$ imply that $x = y$. Prove that if the norm of $X$ is strictly convex, then the point $y_0$ occurring in the assertion of Problem 4.3.2 is unique.

*Solution.* Let $\| \ \|$ be a strictly convex norm on a linear space $X$, and let $Y$ be a finite-dimensional linear subspace of $X$. Let $x_0 \in X - Y$, $L = \inf_{y \in Y} \|x_0 - y\|$, and suppose that there are elements $y_0, y_0' \in Y$ such that

$$\|x_0 - y_0\| = \|x_0 - y_0'\| = L.$$

It is clear that $\|(x_0 - y_0)/L\| = \|(x_0 - y_0')/L\| = 1$. Moreover,

$$\|(x_0 - y_0)/L + (x_0 - y_0')/L\| = \frac{2}{L}\|x_0 - (y_0 + y_0')/2\| \geq \frac{2}{L}L = 2,$$

and

$$\|(x_0 - y_0)/L + (x_0 - y_0')/L\| \leq \frac{1}{L}\left(\|x_0 - y_0\| + \|x_0 - y_0'\|\right) = \frac{1}{L}2L = 2,$$

so $\|(x_0 - y_0)/L + (x_0 - y_0')/L\| = 2$. By strict convexity,

$$\frac{1}{L}(x_0 - y_0) = \frac{1}{L}(x_0 - y_0'),$$

and it follows that $y_0 = y_0'$.

**4.3.4**

Prove that the norm of $L^p(X, \mu)$ is strictly convex if $1 < p < \infty$, and is not strictly convex if $p = 1$ or if $p = \infty$.

*Solution.* Let $f, g \in \mathscr{L}^p(X, \mu)$, $1 < p < \infty$, and suppose $\|f\|_p = \|g\|_p = 1$ and $\|f + g\|_p = 2$. We then have equality in Minkowski's inequality:

$$\|f + g\|_p = \|f\|_p + \|g\|_p.$$

By Problem 3.2.7 this implies that $f = 0$ a.e., or $g = 0$ a.e., or $f = \lambda g$ a.e. for some positive constant $\lambda$. The first two possibilities are ruled out since $\|f\|_p = \|g\|_p = 1$, so the third alternative must hold. But then

$$1 = \|f\|_p = |\lambda| \|g\|_p = |\lambda|,$$

so $\lambda = 1$. Thus $f = g$ a.e., so that $\tilde{f} = \tilde{g}$ in $L^p(X, \mu)$. It follows that $\| \ \|_p$ is strictly convex if $1 < p < \infty$.

For $p = 1, \infty$, surely the problem is supposed to say 'not *necessarily* strictly convex', because we can come up with examples where $\| \ \|_p$ *is* strictly convex, such as an empty measure space $X = \emptyset$, or any space with identically zero measure $\mu = 0$. (Perhaps less trivial examples exist.) Hence we will only

demonstrate that there are examples where $\| \ \|_p$ ($p \in \{1, \infty\}$) is not strictly convex.

For $p = 1$, consider the $L^1$-space $l^1$. The sequences $x = (1, 0, 0, \dots)$ and $y = (0, 1, 0, 0, \dots)$ satisfy $\|x\|_1 = \|y\|_1 = 1$ and $\|x + y\|_1 = 2$, yet $x \neq y$. For $p = \infty$, consider the $L^\infty$-space $l^\infty$. The sequences $x = (1, 0, 0, \dots)$ and $y = (1, 1, 1, \dots)$ satisfy $\|x\|_\infty = \|y\|_\infty = 1$ and $\|x + y\|_\infty = 2$, but $x \neq y$.

### 4.3.5

Prove that in $C[a, b]$ the uniform norm is not equivalent to the $L^p$ norm (for $1 \leq p < \infty$).

*Solution.* For simplicity, let $a = 0$ and $b = 1$. For $n = 1, 2, \dots$, define $f_n : [0, 1] \to \mathbb{R}$ (or $\mathbb{C}$) by $f_n(x) = x^n$. Let $\| \ \|_u$ denote the uniform norm. For all $n$ we have

$$\|f_n\|_u = \max_{0 \leq x \leq 1} |f(x)| = 1$$

and

$$\|f_n\|_p = \left( \int_0^1 x^n dx \right)^{1/p} = (n + 1)^{-1/p}.$$

(We are assuming that the $L^p$ norm is defined with the standard Lebesgue measure.) In particular $\|f_n\|_p \to 0$ as $n \to \infty$, whatever the value of $p$ ($\neq \infty$). It follows that there exists no $\beta > 0$ such that $\|f\|_u \leq \beta \|f\|_p$ for all $f \in C[0, 1]$, and hence that the two norms are not equivalent.

### 4.3.7

Let $n$ be a positive integer, $1 \leq p < \infty$, and let $f(x)$ be a continuous function on $0 \leq x \leq 1$. Then there exists a unique polynomial $Q_n$ of degree $n$ such that for any other polynomial $P_n$ of degree $n$

$$\int_0^1 |f(x) - P_n(x)|^p dx > \int_0^1 |f(x) - Q_n(x)|^p dx.$$

*Solution.* There is an error in the problem statement; it should be polynomials of degree $\leq n$, not *exactly* $n$. To see that the written claim is false, consider the case $f(x) = 0$ and $n = 1$. Then $f$ can be approximated arbitrarily closely by degree 1 polynomials $ax + b$, $a \neq 0$, but no such polynomial will make $\|f - P\|_p$ vanish completely.

Let $\mathscr{P}_n$ denote the set of polynomial functions on $[0, 1]$ with degree $\leq n$. Then $\mathscr{P}_n$ is a linear subspace of the normed linear space $(C[0, 1], \| \ \|_p)$, and is finite-dimensional since it is spanned by the polynomials $1, x, x^2, \dots, x^n$.

If $f \in \mathscr{P}_n$, then $Q = f$ satisfies the claim, since $\|f - P\|_p > 0$ for all $P \neq f$ (else $\| \ \|_p$ would not be a norm). If $f \notin \mathscr{P}_n$, then the conclusion of Problem 4.3.2 tells us that there exists $Q \in \mathscr{P}_n$ such that

$$\inf_{P \in \mathscr{P}_n} \|f - P\| = \|f - Q\|.$$

This $Q$ is unique by Problem 4.3.3 if the norm $\| \ \|_p$ is strictly convex, whereupon the claim follows. By Problem 4.3.4 this is the case for $1 < p < \infty$. In fact the norm is strictly convex even for $p = 1$ over $C[0, 1]$, as is easily verified directly from the definition.

## Section 4.4 − Linear Transformations

### Problems

#### 4.4.2

Let $T$ be an additive operator [that is, $T(x_1 + x_2) = Tx_1 + Tx_2$] from a real normed linear space $X$ into a normed linear space $Y$. If $T$ is continuous, then $T$ is homogeneous [that is, $T(\lambda x) = \lambda Tx$]. [*Hint:* Prove that $T[(m/n)x] = (m/n)Tx$, where $m, n$ are integers.]

*Solution.* Let $x$ be an arbitrary element of $X$. By induction, additivity implies $T(mx) = m\, Tx$ for positive integers $m$. Moreover,

$$T0 = T(0 + 0) = T0 + T0$$

implies that $T0 = 0$, so that

$$0 = T0 = T(mx - mx) = T(mx) + T(-mx),$$

which shows that $T(-mx) = -T(mx) = -m\, Tx$, thus extending the earlier result to nonpositive integers. Finally, if $m$ and $n$ are integers, $n \neq 0$, then

$$n\, T\left(\frac{m}{n}x\right) = T\left(n\frac{m}{n}x\right) = T(mx) = m\, Tx,$$

so that $T[(m/n)x] = (m/n)Tx$, further extending the result to rationals.

Now, let $\lambda \in \mathbb{R}$. There exists a sequence $(\lambda_n)$ in $\mathbb{Q}$ such that $\lambda_n \to \lambda$. Clearly $\lambda_n x \to \lambda x$ in $X$, so $T(\lambda_n x) \to T(\lambda x)$ in $Y$ by continuity of $T$. But, by what we found above,

$$T(\lambda_n x) = \lambda_n\, Tx \to \lambda\, Tx,$$

so $T(\lambda x) = \lambda\, Tx$.

#### 4.4.4

Let $f(x) = \sum_{n=0}^{\infty} a_n z^n$ be an entire complex analytic function. Prove that for every $T \in \mathscr{B}(X)$, $X$ a Banach space, the series $\sum_{n=0}^{\infty} a_n T^n$ $(T^0 = I)$ is absolutely convergent in $\mathscr{B}(X)$. One defines $f(T)$ by

$$f(T) = \sum_{n=0}^{\infty} a_n T^n.$$

*Solution.* Recall that every complex power series converges absolutely on the interior of its disc of convergence (see for example Abel's theorem in Ahlfors' *Complex Analysis*). For $\sum_n a_n z^n$ the disc of convergence is the entire complex plane (since $f$ is entire), so $\sum_n |a_n z^n| < \infty$ for all $z \in \mathbb{C}$.

By Theorem 4.4.4, $\mathscr{B}(X)$ is a Banach space and hence a Banach algebra by Example 2. Thus $\|T^n\| = \|T\|^n$ for all $n \in \mathbb{N}$. It follows that

$$\sum_{n=0}^{\infty} \|a_n T^n\| \leq \sum_{n=0}^{\infty} |a_n| \cdot \|T\|^n < \infty.$$

The series $\sum_{n=0}^{\infty} a_n T^n$ is therefore (uniformly) convergent, and defines an operator $f(T) \in \mathscr{B}(X)$.

*Remark:* If $(S_n)$ is a (uniformly) convergent sequence of operators in $\mathscr{B}(X)$ with limit $S$, then
$$|S_n x - S x| \leq \|S_n - S\| \cdot \|x\| \to 0$$

for all $x \in X$, so that $S_n x \to S x$ (pointwise) as well. In particular,

$$f(T)x = \sum_{n=0}^{\infty} a_n T^n x.$$

### 4.4.5

Let $f(z)$, $g(z)$ be entire complex analytic functions, and let $h(z) = f(z)g(z)$. Prove that for any $T \in \mathscr{B}(X)$, $X$ a Banach space, $h(T) = f(T)g(T)$. In particular we have $e^{\lambda T} e^{\mu T} = e^{(\lambda + \mu)T}$.

*Solution.* We will use Mertens' theorem; see for example Theorem 3.50 in Rudin's *Principles of Mathematical Analysis*, and note that the proof remains valid for any Banach algebra.

Since $f$ and $g$ are entire they can be expressed as everywhere convergent power series

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \quad g(z) = \sum_{n=0}^{\infty} b_n z^n.$$

By Mertens' theorem,

$$h(z) = \sum_{n=0}^{\infty} c_n z^n, \quad c_n = \sum_{k=0}^{n} a_k b_{n-k} z^n.$$

Let $f(T)$, $g(T)$, and $h(T)$ be defined as in the previous problem. Then, again by Mertens' theorem (applied to the Banach algebra $\mathscr{B}(X)$),

$$f(T)g(T) = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k T^k b_{n-k} T^{n-k} = \sum_{n=0}^{\infty} \sum_{k=0}^{n} a_k b_{n-k} T^n = \sum_{n=0}^{\infty} c_n T^n = h(T).$$

The exponential function is known from basic complex analysis to be entire, with power series representation

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

(This is either a theorem or a definition, depending on how the subject is developed.) It is also well-known that $e^{z+w} = e^z e^w$ for all $z, w \in \mathbb{C}$. Hence, given $\lambda, \mu \in \mathbb{C}$, we can apply the above to $f(z) = e^{\lambda z}$, $g(z) = e^{\mu z}$, and $h(z) = e^{(\lambda+\mu)z}$, all clearly entire functions, to obtain

$$e^{(\lambda+\mu)T} = e^{\lambda T} e^{\mu T}$$

for all $T \in \mathscr{B}(X)$.

### 4.4.6

Find the norm of the operator $A \in \mathscr{B}(X)$ given by $(Af)(t) = tf(t)$ $(0 \le t \le 1)$, where (a) $X = C[0,1]$, (b) $X = L^p(0,1)$ and $(1 \le p \le \infty)$.

*Solution.*

(a) For $0 \le t \le 1$ we have $|tf(t)| = |t||f(t)| \le |f(t)|$, hence $\|Af\| \le \|f\|$, and

$$\|A\| = \sup_{f \ne 0} \frac{\|Af\|}{\|f\|} \le 1.$$

   The upper bound $\|Af\|/\|f\| = 1$ is attained with $f$ constant, so $\|A\| = 1$.

(b) Let us first verify that $A \in \mathscr{B}(X)$, i.e. that it is a bounded linear operator $L^p(0,1) \to L^p(0,1)$. Linearity is immediate. If $f \in \mathscr{L}^p(0,1)$, then $|f|^p$ is integrable, so $|tf(t)|^p = |t|^p|f(t)|^p$ is integrable by Corollary 2.10.2, and we see that $A$ does indeed map into $L^p(0,1)$. Finally, since $|tf(t)|^p = |t|^p|f(t)|^p \le |f(t)|^p$ for $0 < t < 1$, we have $\|Af\|_p \le \|f\|_p$, so that $\|A\| \le 1$.
   We will now show that $\|A\| \ge 1$, so that $\|A\| = 1$. The case $p = \infty$ is similar to (a), so we will assume $1 \le p < \infty$. For $n = 1, 2, \ldots$, define simple functions $f_n : (0,1) \to \mathbb{R}$ (or $\mathbb{C}$) by

$$f_n(t) = \begin{cases} 0 & \text{if } 0 < t < 1 - 1/n, \\ n^{1/p} & \text{if } 1 - 1/n \le t < 1. \end{cases}$$

   Then one easily finds that $\|f_n\|_p = 1$ and $\|Af_n\|_p \ge 1 - 1/n$, so that

$$\|A\| \ge \frac{\|Af_n\|_p}{\|f_n\|_p} \ge 1 - \frac{1}{n}$$

   for all $n$. Indeed it follows that $\|A\| \ge 1$.

51

### 4.4.7

A linear operator from a normed linear space $X$ into a normed linear space $Y$ is bounded if and only if it maps bounded sets onto bounded sets.

*Solution.* Let $T : X \to Y$ be a linear operator between normed linear spaces.

Suppose first that $T$ is bounded, and let $A$ be a bounded subset of $X$. Then $\|T\| < \infty$, and there is some $L$ such that $\|x\| \leq L < \infty$ for all $x \in A$. Hence

$$\|Tx\| \leq \|T\| \, \|x\| \leq \|T\| L$$

for all $x \in A$, so $T(A)$ is bounded.

Conversely, suppose that $T$ maps bounded sets onto bounded sets. The set $\{x \in X : \|x\| = 1\}$ is bounded, so there exists $M$ such that $\|Tx\| \leq M < \infty$ whenever $\|x\| = 1$. It follows that

$$\|T\| = \sup_{\|x\|=1} \|Tx\| \leq M,$$

and hence that $T$ is bounded.

### 4.4.8

A linear operator from a normed linear space $X$ into a normed linear space $Y$ is continuous if and only if it maps sequences converging to 0 into bounded sequences.

*Solution.* Let $T : X \to Y$ be a linear operator between normed linear spaces. The claim is immediate if $X$ is the trivial space (containing only the 0 vector), so let us assume that $X$ is nontrivial.

Suppose first that $T$ is continuous, and therefore bounded. Any sequence in $X$ converging to 0 is easily seen to be bounded, hence is mapped to a bounded sequence by the conclusion of the previous problem.

Conversely, suppose $T$ has the property that it maps sequences converging to 0 into bounded sequences. Assume towards a contradiction that $T$ is unbounded. Then it is possible to construct a sequence $(x_n)$ in $X$ such that $\|x_n\| = 1$ and $\|Tx_n\| > n$ for every $n$. The sequence $(x_n/\sqrt{n})$ converges to 0, so $\{T(x_n/\sqrt{n})\}$ is bounded by hypothesis. But

$$\left\| T\left( \frac{x_n}{\sqrt{n}} \right) \right\| = \frac{1}{\sqrt{n}} \|Tx_n\| > \frac{1}{\sqrt{n}} \cdot n = \sqrt{n} \to \infty$$

as $n \to \infty$, yielding a contradiction. We conclude that $T$ must be bounded, and thus continuous.

# Section 4.6 – The Open-Mapping Theorem and the Closed-Graph Theorem

## Problems

### 4.6.1

If $T, S, T^{-1}, S^{-1}$ belong to $\mathscr{B}(X)$, then $(TS)^{-1} \in \mathscr{B}(X)$ and $(TS)^{-1} = S^{-1}T^{-1}$.

*Solution.* Note in particular that $T, S, T^{-1}, S^{-1}$ belonging to $\mathscr{B}(X)$ implies that all of these transformations are bijective, in addition to bounded. (If $T$ is not surjective, then $D_{T^{-1}} = T(X) \neq X$, hence $T^{-1} \notin \mathscr{B}(X)$.)

It is clear that $\|TS\| \leq \|T\| \|S\| < \infty$, hence $TS \in \mathscr{B}(X)$, and it is bijective on account of being a composition of bijections. Thus $(TS)^{-1}$ exists and is equal to $S^{-1}T^{-1}$. It follows that $\|(TS)^{-1}\| \leq \|S^{-1}\| \|T^{-1}\| < \infty$, hence $(TS)^{-1} \in \mathscr{B}(X)$.

### 4.6.2

Let $X$ be a Banach space and let $A \in \mathscr{B}(X)$, $\|A\| < 1$. Prove that $(I + A)^{-1}$ exists and is given by

$$(I + A)^{-1} = \sum_{n=0}^{\infty} (-1)^n A^n,$$

where the series is absolutely convergent [in $\mathscr{B}(X)$]. Show also that

$$\|(I + A)^{-1}\| \leq 1/(1 - \|A\|).$$

*Solution.* We note first that

$$\sum_{n=0}^{\infty} \|(-1)^n A^n\| = \sum_{n=0}^{\infty} \|A^n\| \leq \sum_{n=0}^{\infty} \|A\|^n = \frac{1}{1 - \|A\|},$$

since $\|A\| < 1$. Hence the series $\sum_{n=0}^{\infty} (-1)^n A^n$ is strongly convergent to an operator $B \in \mathscr{B}(X)$, by Theorem 4.5.2. (Recall that $\mathscr{B}(X)$ is a Banach space, hence absolute convergence implies convergence. Convergence in $\mathscr{B}(X)$ is the uniform convergence of operators, and uniform convergence implies strong convergence.)

For all $x \in X$, we have

$$ABx = A\left(\sum_{n=0}^{\infty} (-1)^n A^n x\right) = \sum_{n=0}^{\infty} (-1)^n A^{n+1} x = x - \sum_{n=0}^{\infty} (-1)^n A^n x = (I - B)x,$$

where we have used the continuity of $A$ to interchange limiting processes. Also,

$$BAx = \sum_{n=0}^{\infty} (-1)^n A^{n+1} x = ABx.$$

It follows that

$$B(I + A) = (I + A)B = B + AB = B + I - B = I,$$

so that $B = (I + A)^{-1}$.

Finally,

$$\|(I + A)^{-1}\| \leq \sum_{n=0}^{\infty} \|(-1)^n A^n\| \leq \frac{1}{1 - \|A\|}$$

by what we found earlier.

### 4.6.3

Let $X$ be a Banach space and let $T$ and $T^{-1}$ belong to $\mathscr{B}(X)$. Prove that if $S \in \mathscr{B}(X)$ and $\|S - T\| < 1/\|T^{-1}\|$, then $S^{-1}$ exists and is a bounded operator, and

$$\|S^{-1} - T^{-1}\| < \frac{\|T^{-1}\|}{1 - \|S - T\| \, \|T^{-1}\|}.$$

[*Hint:* $S = [(S - T)T^{-1} + I]T$.]

*Solution.* Note that $ST^{-1} = I + (S - T)T^{-1}$, and that

$$\|(S - T)T^{-1}\| \leq \|S - T\| \|T^{-1}\| < 1.$$

By the previous problem $(ST^{-1})^{-1}$ exists, and

$$\|(ST^{-1})^{-1}\| \leq \frac{1}{1 - \|(S - T)T^{-1}\|} \leq \frac{1}{1 - \|(S - T)\| \, \|T^{-1}\|}.$$

Moreover, since $ST^{-1}(ST^{-1})^{-1} = I$ and

$$T^{-1}(ST^{-1})^{-1}S = T^{-1}(ST^{-1})^{-1}ST^{-1}T = T^{-1}T = I,$$

we have $S^{-1} = T^{-1}(ST^{-1})^{-1}$. Finally,

$$\|S^{-1}\| = \|T^{-1}(ST^{-1})^{-1}\| \leq \|T^{-1}\| \, \|(ST^{-1})^{-1}\| \leq \frac{\|T^{-1}\|}{1 - \|(S - T)\| \, \|T^{-1}\|},$$

and

$$\|S^{-1} - T^{-1}\| = \|S^{-1}(S - T)T^{-1}\| \leq \|S^{-1}\| \, \|S - T\| \, \|T^{-1}\| < \|S^{-1}\|,$$

so that

$$\|S^{-1} - T^{-1}\| < \frac{\|T^{-1}\|}{1 - \|(S - T)\| \, \|T^{-1}\|}.$$

**4.6.4**

Let $X$ and $Y$ be two linear vector spaces. Find necessary and sufficient conditions for a subset $G$ of $X \times Y$ to be the graph of a linear operator from $X$ into $Y$.

*Solution.* We claim that $G$ is the graph of a linear operator from $X$ into $Y$ if and only if

  (i) $G$ is a linear subspace of $X \times Y$.

  (ii) The set $G \cap (\{0\} \times Y)$ is a singleton.

It is clear that these conditions are necessary, so we need only prove that they are sufficient.

Since $G$ is nonempty by (ii), it contains an element $(x, y)$. By (i) it also contains $(0, 0) = 0 \cdot (x, y)$, and it follows that $G \cap (\{0\} \times Y) = \{(0, 0)\}$.

If $(u, v), (u, v') \in G$, then

$$(0, v - v') = (u, v) - (u, v') \in G \cap (\{0\} \times Y)$$

by (i). It follows that $(0, v - v') = (0, 0)$, hence that $v = v'$.

By the above, $G$ is a functional relation on $X \times Y$, so it defines a partial function $T : D \subset X \to Y$, where

$$D = \{x \in X; \; \exists y \in Y \text{ such that } (x, y) \in G\},$$

and $T(x) = y$ if $(x, y) \in G$.

Suppose $x_1, x_2 \in D$, and $\lambda_1, \lambda_2$ are scalars. Then

$$(\lambda_1 x_1 + \lambda_2 x_2, \lambda_1 T(x_1) + \lambda_2 T(x_2)) = \lambda_1(x_1, T(x_1)) + \lambda_2(x_2, T(x_2)) \in G$$

by (i). Hence $\lambda_1 x_1 + \lambda_2 x_2 \in D$, which shows that $D$ is a linear subspace of $X \times Y$, and

$$T(\lambda_1 x_1 + \lambda_2 x_2) = \lambda_1 T(x_1) + \lambda_2 T(x_2),$$

so that $T$ is a linear operator.

Finally, note that $G$ is precisely the graph of $T$. This completes the proof.

**4.6.5**

Let $X$ and $Y$ be Banach spaces and let $T$ be a bounded linear map from $X$ into $Y$. If $T(X)$ is of the second category (in $Y$), then $T(X) = Y$.

*Solution.* We will use the following lemma:

> If $W$ is a linear subspace of a normed space $V$, and if $W$ contains a nonempty open subset of $V$, then $W = V$.

To see that this is true, note that $W$ contains an open ball $B(x_0, r)$. Given any $y \in V$, let

$$x = \frac{r}{2\|y\|} y + x_0.$$

55

Then $x \in B(x_0, r) \subset W$, so that

$$y = \frac{2\|y\|}{r}(x - x_0) \in W$$

by closure under linear combinations.

Now, assume that $T(X)$ is of the second category, and follow the steps of parts (a) and (b) of the proof of Theorem 4.6.1, but with $T(X)$ in place of $Y$. We find that $T(X)$ contains an open ball (see equation (4.6.2)) and hence that $T(X) = Y$ by our lemma.

### 4.6.6

Let $X$ and $Y$ be Banach spaces and let $T$ be a linear map from a linear subspace $D_T$ of $X$ into $Y$. If $D_T$ (in $X$) and the graph of $T$ (in $X \times Y$) are closed, then $T$ is bounded—that is, $\|Tx\| \le K\|x\|$ for all $x \in D_T$ ($K$ constant).

*Solution.* Note first that $D_T$ is complete; this is true for any closed subset of any complete space. Indeed, if $(x_n)$ is a Cauchy sequence in $D_T$, then it is also a Cauchy sequence in $X$, so it converges to a point $x \in X$, and since $D_T$ is closed we have $x \in D_T$. It follows that $D_T$ is a Banach space. Hence we can regard $T$ as a map $D_T \to Y$ and apply Theorem 4.6.4 to conclude that $T$ is continuous, thus bounded by Theorem 4.4.2.

### 4.6.7

Let $X$ be a normed linear space with any one of two norms $\|\ \|_1$, $\|\ \|_2$. If $\|x_n\|_2 \to 0$ implies $\|x_n\|_1 \to 0$, then (4.6.5) holds.

*Solution.* Assume towards a contradiction that (4.6.5) does not hold. Then, for every $n \in \{1, 2, \dots\}$, there exists $x_n \in X$ such that $\|x_n\|_1 > n\|x_n\|_2$. The sequence $y_n = x_n/(n\|x_n\|_2)$ satisfies

$$\|y_n\|_1 = \frac{\|x_n\|_1}{n\|x_n\|_2} > \frac{n\|x_n\|_2}{n\|x_n\|_2} = 1$$

for all $n$. But $\|y_n\|_2 = 1/n \to 0$, which by hypothesis implies $\|y_n\|_1 \to 0$, yielding a contradiction.

## Section 4.8 – The Hahn-Banach Theorem

### Problems

### 4.8.1

Let $X$ be a normed linear space and let $\{x_n\} \subset X$. A point $y_0$ is the limit of linear combinations $\sum_{j=1}^{n} c_j x_j$ if and only if $x^*(y_0) = 0$ for all $x^*$ for which $x^*(x_j) = 0$ for $1 \le j < \infty$.

*Solution.* Note that "$y_0$ is the limit of linear combinations $\sum_{j=1}^{n} c_j x_j$" is not meant to imply $y_0 = \lim_{n \to \infty} \sum_{j=1}^{n} c_j x_j$ for some sequence of coefficients $(c_j)$. Rather, it just says that $y_0$ is the limit of a sequence of finite linear combinations of elements in $\{x_n\}$; i.e., that $y_0$ is a limit point of $S := \operatorname{span}\{x_n\}$. Also, the statement "$x^*(x_j) = 0$ for $1 \leq j < \infty$" can be simplified to "$x^*$ vanishes on $S$."

Suppose that $y_0$ is a limit point of $S$. Then there is a sequence $(s_n)$ in $S$ such that $s_n \to y_0$. If $x^* \in X^*$ vanishes on $S$, then

$$x^*(y_0) = x^* \left( \lim_{n \to \infty} s_n \right) = \lim_{n \to \infty} x^*(s_n) = 0$$

by continuity of $x^*$.

Conversely, suppose that $x^*(y_0) = 0$ for all $x^* \in X^*$ that vanish on $S$. Let

$$d = \inf_{s \in S} \|s - y_0\|.$$

If $d > 0$, then Theorem 4.8.3 tells us that there exists $x^* \in X^*$ such that $x^*(y_0) = 1$, but which vanishes on $S$. This contradicts our assumptions, so we conclude that $d = 0$, which in turn shows that $y_0$ is a limit point of $S$.

### 4.8.5

Let $X$ be an infinite-dimensional Banach space. Prove that there exists an infinite, strictly decreasing sequence $\{Y_n\}$ of infinite-dimensional closed linear subspaces of $X$. [*Hint:* Take $Y_1$ to be the null space of some $x_1^* \neq 0$ in $X^*$. Take $Y_2$ to be the null space of some $x_2^* \neq 0$ in $Y_1^*$, and so on.]

*Solution.* We will construct such a sequence by induction. For the base case, let $Y_0 = X$. Certainly $Y_0$ is an infinite-dimensional closed linear subspace of $X$.

For the inductive step, assume that we have infinite-dimensional closed linear subspaces $Y_0 \supset Y_1 \supset \cdots \supset Y_n$, with the inclusions being strict. Fix an element $x_0 \in Y_n$ such that $\|x_0\| = 1$. (Such an element is guaranteed to exist since $Y_n$ is infinite-dimensional.) Corollary 4.8.4 provides a continuous linear functional $x^* \in Y_n$ such that $x^*(x_0) = 1$. Let $Y_{n+1}$ be the null space of $x^*$; clearly a proper linear subspace of $Y_n$, hence also of $X$. As discussed after Corollary 4.8.7, each element $x \in Y_n$ can be written as $x = z + \lambda x_0$, where $\lambda = x^*(x)$ and $z = x - \lambda x_0 \in Y_{n+1}$. It follows that $Y_{n+1}$ is infinite-dimensional. Finally, if $(y_i)$ is a sequence in $Y_{n+1}$ and $y_i \to y \in X$, then $x^*(y) = \lim_{i \to \infty} x^*(y_i) = 0$ by continuity, so $y \in Y_{n+1}$, making $Y_{n+1}$ a closed subset of $X$. This concludes the inductive step.

### 4.8.9

Let $u(t)$ be a function defined on $a < t < b$ with values in a Banach space $X$. We say that $u(t)$ is *strongly differentiable at $t$ [on $(a,b)$]* if $\lim_{h \to 0}\{[u(t + h) - u(t)]/h\}$ exists [for all $t \in (a,b)$]. The limit is denoted by $du(t)/dt$ and is called the derivative of $u(t)$. For functions $A(t)$ with values in $\mathscr{B}(X)$, if $\lim_{h \to 0}\{[A(t + h)x - A(t)x]/h\}$ exists for any $x \in X$, then we say that $A(t)$

has a *strong derivative*. If $\lim_{h\to 0}\{[A(t + h) - A(t)]/h\}$ exists (in the uniform topology), then we say that $A(t)$ is *uniformly differentiable*. Prove that $e^{tA}$ $[A \in \mathscr{B}(X)]$ is uniformly differentiable and $de^{tA}/dt = Ae^{tA}$.

*Solution.* Fix any $t \in \mathbb{R}$ and $A \in \mathscr{B}(X)$. By Problem 4.4.4,

$$e^{tA} = \sum_{n=0}^{\infty} \frac{(tA)^n}{n!} \in \mathscr{B}(X),$$

and by Problem 4.4.5 we have $e^{(t+h)A} = e^{hA}e^{tA}$ for all $h \in \mathbb{R}$. It follows that

$$\left\| \frac{e^{(t+h)A} - e^{tA}}{h} - Ae^{tA} \right\| = \left\| \left( \frac{e^{hA} - I}{h} - A \right) e^{tA} \right\| \le \left\| \frac{e^{hA} - I}{h} - A \right\| \cdot \left\| e^{tA} \right\|,$$

hence it is sufficient to show that

$$\lim_{h\to 0} \frac{e^{hA} - I}{h} = A.$$

To that end, note that

$$\frac{e^{hA} - I}{h} - A = \frac{1}{h} \left( \sum_{n=0}^{\infty} \frac{(hA)^n}{n!} - I - hA \right) = \frac{1}{h} \sum_{n=2}^{\infty} \frac{(hA)^n}{n!}.$$

Thus

$$\left\| \frac{e^{hA} - I}{h} - A \right\| \le \frac{1}{|h|} \sum_{n=2}^{\infty} \frac{|h|^n \cdot \|A\|^n}{n!}$$

$$= |h| \cdot \|A\|^2 \sum_{n=2}^{\infty} \frac{|h|^{n-2} \cdot \|A\|^{n-2}}{n!}$$

$$= |h| \cdot \|A\|^2 \sum_{n=0}^{\infty} \frac{(|h| \cdot \|A\|)^n}{(n + 2)!}$$

$$\le |h| \cdot \|A\|^2 \cdot e^{|h| \cdot \|A\|},$$

which goes to zero as $h \to 0$. It follows that $e^{tA}$ is uniformly differentable on $\mathbb{R}$, with derivative $Ae^{tA}$.

### 4.8.10

Let $X$ be a real normed linear space, and let $u(t)$ be continuous and strongly differentiable in $(a, b)$. Then for any $a < \alpha < \beta < b$,

$$\|u(\beta) - u(\alpha)\| \le (\beta - \alpha) \sup_{\alpha \le t \le \beta} \left\| \frac{du(t)}{dt} \right\|.$$

[*Hint:* Apply $x^*$ to $u(\beta) - u(\alpha)$.]

*Solution.* If $u(\alpha) = u(\beta)$ then there is nothing to prove, so assume $u(\beta) \neq u(\alpha)$. By Corollary 4.8.4 there is a bounded linear operator $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(u(\beta) - u(\alpha)) = \|u(\beta) - u(\alpha)\|$. Define $f = x^* \circ u : (a, b) \to \mathbb{R}$. Since $x^*$ is linear and continuous,

$$\lim_{h \to 0} \frac{f(t+h) - f(t)}{h} = \lim_{h \to 0} x^* \left[ \frac{u(t+h) - u(h)}{h} \right] = x^* \left[ \lim_{h \to 0} \frac{u(t+h) - u(h)}{h} \right],$$

so that $f$ is differentiable with derivative

$$\frac{df(t)}{dt} = x^* \left[ \frac{du(t)}{dt} \right].$$

By the mean value theorem from elementary real analysis, we have

$$f(\beta) - f(\alpha) = (\beta - \alpha) \frac{df(\gamma)}{dt}$$

for some $\gamma \in (\alpha, \beta)$. Thus,

$$
\begin{aligned}
\|u(\beta) - u(\alpha)\| &= x^*(u(\beta) - u(\alpha)) \\
&= f(\beta) - f(\alpha) \\
&= (\beta - \alpha) \frac{df(\gamma)}{dt} \\
&= (\beta - \alpha) x^* \left[ \frac{du(\gamma)}{dt} \right] \\
&\leq (\beta - \alpha) \|x^*\| \cdot \left\| \frac{du(\gamma)}{dt} \right\| \\
&\leq (\beta - \alpha) \sup_{\alpha \leq t \leq \beta} \left\| \frac{du(t)}{dt} \right\|.
\end{aligned}
$$

**4.8.12**

For every normed linear space $X$ there is a set $A$ such that $X$ is isomorphic to a subspace of the Banach space of functions $f$ on $A$ with norm $\|f\| = \sup_{t \in A} |f(t)|$. If $X$ is separable, $A$ is countable. [*Hint:* Let $\{x_\alpha : \alpha \in A\}$ be dense in $X$. Let $f(x, \alpha)$ be the bounded linear functional (in $x \in X$) satisfying $\|f(\,.\,, \alpha)\| = 1$, $f(x_\alpha, \alpha) = \|x_\alpha\|$. Define the isomorphism $x \to g_x(\alpha)$, where $g_x(\alpha) = f(x, \alpha)$. Prove: $\big| |f(x, \alpha)| - \|x\| \big| \leq 2\|x_\alpha - x\|.$]

*Solution.* The conclusion is immediate if $X$ is a trivial space, so assume $X \neq \{0\}$. Let $\{x_\alpha\}_{\alpha \in A}$ be any dense subset of $X$ (possibly $X$ itself), indexed by a set $A$. (Recall that any space can be indexed by itself.) For each $\alpha \in A$, Corollary 4.8.4 fashions a bounded linear functional $f_\alpha \in X^*$ such that $\|f_\alpha\| = 1$ and $f_\alpha(x_\alpha) = \|x_\alpha\|$.

For each $x \in X$, define $g_x : A \to \mathbb{F}$ (where $\mathbb{F}$ is the field associated with $X$) by $g_x(\alpha) = f_\alpha(x)$. We will prove the following properties of the functions $g_x$:

59

(i) If $x, y \in X$ and $\lambda, \mu \in \mathbb{F}$, then $\lambda g_x + \mu g_y = g_{\lambda x + \mu y}$.

(ii) Each function $g_x$ is bounded, with $\sup_{\alpha \in A} |g_x(\alpha)| = \|x\|$.

To prove (i), simply let $\alpha \in A$ and compute

$$(\lambda g_x + \mu g_y)(\alpha) = \lambda g_x(\alpha) + \mu g_y(\alpha) = \lambda f_\alpha(x) + \mu f_\alpha(y) = f_\alpha(\lambda x + \mu y) = g_{\lambda x + \mu y}(\alpha),$$

using the linearity of $f_\alpha$.

Proving (ii) takes more work. Note first that

$$\sup_{\alpha \in A} |g_x(\alpha)| = \sup_{\alpha \in A} |f_\alpha(x)| \le \sup_{\alpha \in A} \|f_\alpha\| \, \|x\| = \|x\|.$$

Next, given any $\epsilon > 0$, fix $\alpha' \in A$ such that $\|x_{\alpha'} - x\| < \epsilon/2$. (This is possible because $\{x_\alpha\}_{\alpha \in A}$ is dense in $X$.) By the triangle inequality,

$$\|x\| \le \big|\|x\| - \|x_{\alpha'}\|\big| + \big|\|x_{\alpha'}\| - |f_{\alpha'}(x)|\big| + |f_{\alpha'}(x)|.$$

Now, the "reverse triangle inequality" yields

$$\big|\|x\| - \|x_{\alpha'}\|\big| \le \|x_{\alpha'} - x\| < \frac{\epsilon}{2}$$

and

$$
\begin{aligned}
\big|\|x_{\alpha'}\| - |f_{\alpha'}(x)|\big| &\le \big|\|x_{\alpha'}\| - f_{\alpha'}(x)\big| \\
&= |f_{\alpha'}(x_{\alpha'}) - f_{\alpha'}(x)| \\
&= |f_{\alpha'}(x_{\alpha'} - x)| \\
&\le \|f_{\alpha'}\| \, \|x_{\alpha'} - x\| \\
&= \|x_{\alpha'} - x\| \\
&< \frac{\epsilon}{2}.
\end{aligned}
$$

Also,

$$|f_{\alpha'}(x)| \le \sup_{\alpha \in A} |f_\alpha(x)| = \sup_{\alpha \in A} |g_x(\alpha)|.$$

Putting it all together, we have

$$\|x\| \le \sup_{\alpha \in A} |g_x(\alpha)| + \epsilon,$$

and since this holds for every $\epsilon$ we finally arrive at (ii).

Let $F_A$ be the Banach space of all bounded functions $A \to \mathbb{F}$, with the supremum norm $\|g\| = \sup_{\alpha \in A} |g(\alpha)|$. Since the functions $g_x$ are bounded by (ii), we can define a map $\sigma : X \to F_A$ by $\sigma(x) = g_x$. The image of $\sigma$ is a linear subspace of $F_A$ by (i). Moreover, $\sigma$ is an imbedding, since

$$\|\sigma(x) - \sigma(y)\| = \|g_x - g_y\| = \|g_{x-y}\| = \sup_{\alpha \in A} |g_{x-y}(\alpha)| = \|x - y\|$$

for all $x, y \in X$, by (i) and (ii). Every imbedding injective, so $\sigma$ is an isomorphism (in the sense of Section 3.3) onto its image.

Finally, if $X$ is separable, then we can take $A$ to be countable.

# Section 4.10 – Conjugate Spaces and Reflexive Spaces

## Problems

### 4.10.1

A sequence $\{x_n\}$ cannot have two distinct weak limits.

*Solution.* Let $(x_n)$ be a sequence in a normed linear space $X$, and suppose $(x_n)$ converges weakly to both $x$ and $y$ in $X$. Then

$$x^*(x - y) = x^*(x) - x^*(y) = \lim_n x^*(x_n) - \lim_n x^*(x_n) = 0$$

for all $x^* \in X^*$, so $x - y = 0$ by Corollary 4.8.4.

### 4.10.2

Let $X$ be a normed linear space and let $B$ be a dense subset of $X^*$. If a sequence $\{x_n\}$ in $X$ is bounded, and if $\lim_n x^*(x_n)$ exists for each $x^* \in B$, then $\lim_n x^*(x_n)$ exists for all $x^*$ in $X^*$.

*Solution.* Let $X$, $B$, and $(x_n)$ be as in the problem description. Fix any $x^* \in X^*$; we will show that $(x^*(x_n))$ is a Cauchy sequence (in $\mathbb{R}$ or $\mathbb{C}$), and hence that $\lim_n x^*(x_n)$ exists.

Given $\epsilon > 0$, choose $y^* \in B$ such that $\|x^* - y^*\| < \epsilon/4M$, where $M > 0$ is an upper bound of $\{\|x_n\|\}$. Since $\lim_n y^*(x_n)$ exists, there is some $N \in \mathbb{N}$ such that

$$m, n \geq N \implies \|y^*(x_m) - y^*(x_n)\| < \epsilon/2.$$

It follows that

$$\begin{aligned}
\|x^*(x_m) - x^*(x_n)\| &\leq \|x^*(x_m) - y^*(x_m)\| + \|y^*(x_m) - y^*(x_n)\| \\
&\quad + \|y^*(x_n) - x^*(x_n)\| \\
&< \|x^* - y^*\| \cdot \|x_m\| + \frac{\epsilon}{2} + \|y^* - x^*\| \cdot \|x_n\| \\
&< \epsilon
\end{aligned}$$

for all $m, n \geq N$, and hence that $(x^*(x_n))$ is a Cauchy sequence.

### 4.10.3

In a finite-dimensional normed linear space, the concepts of convergence and weak convergene coincide.

*Solution.* It is clear that (strong) convergence implies weak convergence (also for infinite-dimensional spaces), hence we need only prove the converse. Moreover, since any two norms on a finite-dimensional linear space are equivalent by Problem 4.3.1, it suffices to prove the result for any one particular norm.

Let $X$ be a finite-dimensional linear space, with $\{e_1, \ldots, e_n\}$ a basis. Define a norm $\|\ \|$ on $X$ by

$$\left\| \sum_{i=1}^{n} \lambda_i e_i \right\| = \sum_{i=1}^{n} |\lambda_i|.$$

Next, define $e_1^*, \ldots, e_n^* \in X^*$ by

$$e_i^*(e_j) = \begin{cases} 1 & \text{if } i = j, \\ 0 & \text{if } i \neq j. \end{cases}$$

It is easily verified that the functionals $e_i^*$ are linearly independent, and that every $x^* \in X^*$ can be written

$$x^* = \sum_{i=1}^{n} x^*(e_i) e_i^*.$$

Hence $\{e_1^*, \ldots, e_n^*\}$ is a basis of $X^*$. Note also that

$$e_j^* \left( \sum_{i=1}^{n} \lambda_i e_i \right) = \lambda_j,$$

so in fact

$$x = \sum_{i=1}^{n} e_i^*(x) e_i$$

for all $x \in X$.

Now let $(x_k)$ be a weakly convergent sequence in $X$, with weak limit $x$. Then

$$\|x_k - x\| = \left\| \sum_{i=1}^{n} e_i^*(x_k - x) e_i \right\| = \sum_{i=1}^{n} |e_i^*(x_k) - e_i^*(x)|.$$

The right-hand side of this inequality goes to 0 as $k \to \infty$ by weak convergence, hence $\|x_k - x\| \to 0$ as well, and $(x_k)$ converges (strongly) to $x$.

### 4.10.7

A normed linear space that is weakly complete is complete.

*Solution.* Let $(x_n)$ be a Cauchy sequence in a weakly complete normed linear space $X$. For all $x^* \in X^*$ we have

$$|x^*(x_m) - x^*(x_n)| \leq \|x^*\| \cdot \|x_m - x_n\| \to 0 \quad \text{as } m, n \to \infty,$$

so $(x_n)$ is weakly Cauchy as well. Thus $(x_n)$ has a weak limit $x \in X$, by the weak completeness of $X$.

Assume towards a contradiction that $(x_n)$ does not converge (strongly) to $x$. Then there exists $\epsilon > 0$ such that $\|x_n - x\| \geq \epsilon$ for infinitely many $n$. Let

$(y_n)$ be a subsequence such that $\|y_n - x\| \geq \epsilon$ and $\|y_m - y_n\| < \epsilon/2$ for all $m, n \in \mathbb{N}$. By Corollary 4.8.4 there exists $x^* \in X^*$ such that $\|x^*\| = 1$ and $x^*(y_1 - x) = \|y_1 - x\|$, hence we have

$$
\begin{aligned}
x^*(y_n) - x^*(x) &= x^*(y_1 - x) - x^*(y_1 - y_n) \\
&\geq \|y_1 - x\| - \|x^*\| \cdot \|y_1 - y_n\| \\
&> \epsilon/2
\end{aligned}
$$

for all $n$. But $|x^*(y_n) - x^*(x)| \to 0$ since $(y_n)$ is a subsequence of $(x_n)$, so we have a contradiction. Thus $(x_n)$ converges (strongly) to $x$, and $X$ is complete.

# Section 4.12 – Weak Topology in Conjugate Spaces

## Notes

### Terminology: Weak and Weak-*

The topology introduced in this section is called the *weak-\* topology* by other authors, and the convergence defined in Definition 4.12.2 is called *weak-\* convergence*. Indeed this is a better name, since we could also consider weak convergence in $X^*$ in the sense of Definition 4.10.2; i.e. that $x_n^* \to x^*$ weakly if $\lim_n x^{**}(x_n^*) = x^{**}(x^*)$ for all $x^{**} \in X^{**}$. The two types of convergence are equivalent if $X$ is reflexive, but not in general, hence using the same term is problematic.

## Problems

### 4.12.1

Let $X$ be a normed linear space. A sequence $\{f_n\}$ in $X^*$ is weakly convergent to $f \in X^*$ if and only if the following conditions hold: (i) the sequence $\{\|f_n\|\}$ is bounded, and (ii) $\lim_n f_n(x) = f(x)$ for all $x$ in a dense subset of $X$.

*Solution.* There is an error in the problem statement; $(f_n)$ being weakly convergent to $f$ does not in general imply that $\{\|f_n\|\}$ is bounded, unless $X$ is a Banach space (in which case it follows by the uniform boundedness principle). For a counterexample, let $X$ be the subspace of $l^\infty$ consisting of the sequences with only finitely many nonzero entries. Define linear functionals $x_n^* \in X^*$ by $x_n^*(x) = nx_n$ (where $x_n$ denotes the $n$th entry of the sequence $x$). It is easily seen that $\|x_n^*\| = n$, hence $\{\|x_n^*\|\}$ is unbounded. On the other hand $\lim_n x_n^*(x) = 0$ for all $x \in X$, since every $x$ is eventually zero, so $(x_n^*)$ is weakly convergent to the zero functional.

The converse direction does hold as stated; if (i) and (ii) hold, then $(f_n)$ is weakly convergent to $f$ (even if $X$ is not Banach). To prove this, let $A \subset X$ be

the dense subset in (ii), and let $K < \infty$ be an upper bound of $\{\|f_n\|\}$. Fix any $x \in X$. Given $\epsilon > 0$, there exists $y \in A$ such that

$$\|x - y\| < \frac{\epsilon}{2(K + \|f\|)},$$

and $N \in \mathbb{N}$ such that

$$n \geq N \implies |f_n(y) - f(y)| < \frac{\epsilon}{2}.$$

Thus

$$
\begin{aligned}
|f_n(x) - f(x)| &\leq |f_n(x) - f_n(y)| + |f_n(y) - f(y)| + |f(y) - f(x)| \\
&= \|f_n\| \cdot \|x - y\| + \frac{\epsilon}{2} + \|f\| \cdot \|x - y\| \\
&\leq (K + \|f\|) \cdot \|x - y\| + \frac{\epsilon}{2} \\
&< \epsilon
\end{aligned}
$$

for sufficiently large $n$. It follows that $f_n(x) \to f(x)$, and hence that $(f_n)$ is weakly convergent to $f$, since $x$ was arbitrary.

### 4.12.3

Give a direct proof of Theorem 4.12.3. [*Hint:* Choose a subsequence $\{x_{n,1}^*\}$ such that $\lim x_{n,1}^*(x_1)$ exists. Choose a subsequence $\{x_{n,2}^*\}$ such that $\lim x_{n,2}^*(x_2)$ exists, and so on. Then $\{x_{n,n}^*\}$ is weakly convergent.]

*Solution.* Let $X$ be a separable normed linear space over a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, and let $(x_n^*)$ be a bounded sequence in $X^*$. We will show that $(x_n^*)$ has a weakly convergent subsequence.

Fix a countable dense subset $A = \{x_n\} \subset X$. Note that $(x_n^*(x_1))$ is a bounded sequence in $\mathbb{F}$, hence it has a convergent subsequence by the Bolzano-Weierstrass theorem of elementary real analysis. That is, there is a subsequence $(x_{n,1}^*)$ of $(x_n^*)$ such that $(x_{n,1}^*(x_1))$ converges. Similarly $(x_{n,1}^*)$ has a subsequence $(x_{n,2}^*)$ such that $(x_{n,2}^*(x_2))$ converges, and so on.

Consider the sequence $(x_{n,n}^*)$. For each $k \in \mathbb{N}$, excluding the first $k-1$ terms makes it a subsequence of $(x_{n,k}^*)$, so that $\lim_n x_{n,n}^*(x_k) = \lim_n x_{n,k}^*(x_k)$. That is, the sequence $(x_{n,n}^*(x))$ converges for every $x \in A$. In fact $(x_{n,n}^*(y))$ converges for all $y$ in the linear span $Y$ of $A$; if $y = \sum_{i=1}^m \lambda_i x_{k_i}$, then

$$\lim_n x_{n,n}^*(y) = \lim_n x_{n,n}^* \left( \sum_{i=1}^m \lambda_i x_{k_i} \right) = \sum_{i=1}^m \lambda_i \lim_n x_{n,n}^*(x_{k_i}).$$

Hence define $f : Y \to \mathbb{F}$ by $f(y) = \lim_n x_{n,n}^*(y)$. It is clear that $f$ is linear and bounded, so $f \in Y^*$. By Theorem 4.8.2, we can extend $f$ to a bounded linear functional $F \in X^*$. We then have $\lim_n x_{n,n}^*(x) = F(x)$ for all $x \in A$ by construction, thus $(x_{n,n}^*)$ is weakly convergent to $F$ by Problem 4.12.1.

# Section 4.13 – Adjoint Operators

## Problems

### 4.13.2

Let $f(z)$ be an entire complex analytic function and let $T \in \mathscr{B}(X)$. Prove that $[f(T)]^* = f(T^*)$.

*Solution.* Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$. Note that $(T^n)^* = (T^*)^n$ for $n \in \mathbb{N}$ by Theorem 4.13.2. Hence, for all $y^* \in Y^*$ and $x \in X$,

$$
\begin{aligned}
([f(T)]^* y^*)(x) &= y^*(f(T)x) \\
&= y^* \left( \sum_{n=0}^{\infty} a_n T^n x \right) \\
&= \sum_{n=0}^{\infty} a_n y^*(T^n x) \\
&= \sum_{n=0}^{\infty} a_n ((T^n)^* y^*)(x) \\
&= \sum_{n=0}^{\infty} a_n ((T^*)^n y^*)(x) \\
&= \left( \sum_{n=0}^{\infty} a_n (T^*)^n y^* \right)(x) \\
&= (f(T^*) y^*)(x),
\end{aligned}
$$

where we have used the continuity of $y^*$ to interchange limiting processes. The result follows.

### 4.13.3

Let $X$ and $Y$ be Banach spaces and let $T \in \mathscr{B}(X, Y)$. Prove:

(a) If $T$ has a continuous inverse, then $R_T$ is closed.

(b) $T^*$ is one-to-one if and only if $R_T$ is dense in $Y$.

(c) If $T$ maps $X$ onto $Y$, then $T^*$ has a bounded inverse with domain $R_{T^*}$.

*Solution.*

(a) Let $(y_n)$ be a sequence in $R_T$ with limit $y \in Y$, and let $x_n = T^{-1} y_n$ for each $n \in \mathbb{N}$. Note that $(y_n)$ is Cauchy, and it follows (since $T^{-1}$ is bounded) that $(x_n)$ is Cauchy as well;

$$
\|x_m - x_n\| = \|T^{-1} y_m - T^{-1} y_n\| \leq \|T^{-1}\| \cdot \|y_m - y_n\|.
$$

Hence $(x_n)$ has a limit $x \in X$, and we have

$$Tx = T(\lim_n x_n) = \lim_n Tx_n = \lim_n y_n = y.$$

Thus $y \in R_T$, and we conclude that $R_T$ is closed.

(b) We prove the forward implication in the contrapositive; if $R_T$ is not dense in $Y$, then $T^*$ is not injective. Indeed, if $R_T$ is not dense, then Corollary 4.8.7 implies that there exists $y^* \in Y^*$, $y^* \neq 0$, such that $y^*(y) = 0$ for all $y \in R_T$. But then $(T^*y^*)(x) = y^*(Tx) = 0$ for all $x \in X$, so $T^*$ has nontrivial kernel and is therefore not injective.

For the converse, suppose that $R_T$ is dense in $Y$, and let $y_1^*, y_2^* \in Y^*$ be such that $T^*y_1^* = T^*y_2^*$. If $y \in Y$, then there exists a sequence $(Tx_n)$ in $R_T$ such that $Tx_n \to y$, and we have

$$y_1^*(y) = \lim_n y_1^*(Tx_n) = \lim_n y_2^*(Tx_n) = y_2^*(y)$$

by continuity. Thus $y_1^* = y_2^*$, and $T^*$ is injective.

(c) It follows by (b) that $T^*$ is injective. Moreover, its range $R_{T^*}$ is closed by Theorem 4.13.6, and therefore a Banach space. Hence we can regard $T^*$ as an operator $Y^* \to R_{T^*}$ and apply Theorem 4.6.2 to find that $(T^{-1})^* : R_{T^*} \to Y^*$ is bounded.

# Chapter 5 – Completely Continuous Operators

## Section 5.1 – Basic Properties

### Notes

**Characterization in terms of sequences**

Let $X$ and $Y$ be normed linear spaces. An operator $T \in \mathscr{B}(X, Y)$ is compact if and only if it has the following property: if $(x_n)$ is a bounded sequence in $X$, then $(Tx_n)$ has a convergent subsequence.

*Proof.* Suppose first that $T$ is compact. If $(x_n)$ is a bounded sequence in $X$, then $\{Tx_n\} \subset K$ for some compact subset $K \subset Y$. But $K$ is sequentially compact by Theorem 3.5.4, hence $(Tx_n)$ has a convergent subsequence.

   Conversely, suppose $T$ has the property that if $(x_n)$ is a bounded sequence in $X$, then $(Tx_n)$ has a convergent subsequence. Let $A \subset X$ be bounded. Every sequence in $T(A)$ is the image of some (bounded) sequence in $A$, and therefore has a convergent subsequence. It follows that $\overline{T(A)}$ is sequentially compact, and thus compact by Theorem 3.5.4. $\qquad\square$

**Proof Theorem 5.1.3**

Let $A \subset X$ be bounded, with $\|x\| < K < \infty$ for all $x \in X$. Then $\|Sx\| \leq \|S\| \cdot K$ for all $x \in A$, so $S(A)$ is bounded. It follows that $(TS)(A) = T(S(A))$ is relatively compact since $T$ is compact, and hence that $TS$ is compact. Moreover, $S(T(A)) \subset S(\overline{T(A)})$, and the latter set is compact since it is a continuous image of the compact set $\overline{T(A)}$. Thus $ST$ is compact as well.

### Problems

**5.1.1**

A continuous linear transformation with finite-dimensional range is compact.

*Solution.* Let $X$ and $Y$ be normed linear spaces, and let $T \in \mathscr{B}(X, Y)$ with $T(X)$ finite-dimensional. Note that $T(X)$ is closed by Theorem 4.3.2. Let $A \subset X$ be bounded; $\|x\| < K < \infty$ for all $x \in A$. Then $\|Tx\| \leq \|T\| \cdot K$ for all $x \in A$, so $T(A)$ is bounded as well. It follows by Theorem 4.3.3 that $T(A)$ is relatively compact in $T(X)$. Since $T(X)$ is closed, the closure of $T(A)$ in $T(X)$ is also its closure in $Y$, hence $\overline{T(A)}$ is compact in $Y$ as well.

### 5.1.2

A linear combination of completely continuous linear transformations is completely continuous.

*Solution.* It suffices to show that any linear combination of just two compact operators is compact; the claim then follows by induction. Hence let $X$ and $Y$ be normed linear spaces, let $T, S \in \mathscr{B}(X, Y)$ be compact, and let $\lambda, \mu$ be scalars. Let $(x_n)$ be a bounded sequence in $X$. Then $(Tx_n)$ has a convergent subsequence; i.e. there is a subsequence $(y_n)$ of $(x_n)$ such that $(Ty_n)$ converges. Likewise, there is a subsequence $(z_n)$ of $(y_n)$ such that $(Sz_n)$ converges. It follows that the sequence $(\lambda Tz_n + \mu Sz_n)$ is convergent. Hence $\lambda T + \mu S$ is a compact operator.

### 5.1.5

Is the operator $T$ defined by $(Tx)(t) = tx(t)$ $(0 < t < 1)$ completely continuous in $L^2(0, 1)$?

*Solution.* We start with a small generalization of our notion of Riemann integral. Given a bounded and a.e. continuous function $f : (a, b) \to \mathbb{F}$ (where $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$), let $\tilde{f} : [a, b] \to \mathbb{F}$ be an extension of $f$. Then $\tilde{f}$ is Riemann-integrable by Theorem 2.11.1. Moreover, by the elementary theory of the Riemann integral, the values of $\tilde{f}(a)$ and $\tilde{f}(b)$ do not affect the value of the integral. Hence we *define* the Riemann integral of $f$ to be the Riemann-integral of *any* such extension $\tilde{f}$. Note that a result analogous to Theorem 2.11.2 holds for $f$;

$$\int_{(a,b)} f(x)\,dx = \int_{[a,b]} \tilde{f}(x)\,dx = \int_a^b \tilde{f}(x)\,dx = \int_a^b f(x)\,dx,$$

with the second equality following from Theorem 2.11.2, and the last one being our new definition of the Riemann-integral of $f$.

Define functions $x_n : (0, 1) \to \mathbb{R}$ by

$$x_n(t) = \begin{cases} 0 & \text{if } t < 1/2, \\ t^{-1} \sin(2\pi nt) & \text{if } t \geq 1/2. \end{cases}$$

By the preceding discussion each $x_n$ is Riemann-integrable, and also Lebesgue-integrable, with the two integrals agreeing. The same goes for $|x_n|^2$, $|Tx_n|^2$, and $|Tx_n - Tx_m|^2$. We have

$$\int_0^1 |x_n(x)|^2 dx \leq \int_0^1 2\,dx = 2,$$

so $(x_n)$ is a bounded sequence in $L^2(0,1)$. By a rather tedious computation of Riemann-integrals, one finds that $\|Tx_m - Tx_n\|_2 = 1/2$ whenever $m \neq n$, hence no subsequence of $(Tx_n)$ is Cauchy. Since $L^2(0,1)$ is a Banach space, this implies that every subsequence is divergent, and thus that $T$ is not a compact operator.

### 5.1.6

Let $f(z)$ be an entire complex analytic function with $f(0) = 0$, and let $X$ be a Banach space. If $T \in \mathscr{B}(X)$ is completely continuous, then $f(T)$ is also completely continuous.

*Solution.* Since $f$ is entire and $f(0) = 0$, it has a power series representation $f(z) = \sum_{n=1}^{\infty} a_n z^n$. By Theorem 5.1.3, $T^n$ is compact for every $n \geq 1$, hence it follows by Problem 5.1.2 that $\sum_{n=1}^{k} a_n T^n$ is compact for every $k \geq 1$. Since $f(T)$ is the uniform limit of the operators $\sum_{n=1}^{k} a_n T^n$ (c.f. Problem 4.4.4), Theorem 5.1.2 then implies that $f(T)$ is compact.

### 5.1.7

Let $T$ be a compact linear operator from a Banach space $X$ onto itself. If $T^{-1}$ is a bounded operator, then $X$ is finite-dimensional.

*Solution.* The assumption that $T$ is onto is unnecessary; we need only assume that $T$ is injective. By Theorem 5.1.3, the identity operator $I = T^{-1}T$ is compact, hence every bounded subset of $X$ is relatively compact. It follows by Theorem 4.3.3 that $X$ is finite-dimensional.

## Section 5.2 – The Fredholm-Riesz-Schauder Theory

### Problems

### 5.2.4

Consider the *Volterra integral equation*

$$f(s) = g(s) + \int_0^t K(s,t)f(t)\,dt \quad (0 \leq t \leq 1), \tag{5.2.6}$$

where $K(s,t)$ is continuous for $0 \leq s,t \leq 1$. Prove that for any continuous function $g$ there exists a unique continuous solution $f$ of (5.2.6).

*Solution.* Note that there is an error in (5.2.6). It should read (something like)

$$f(t) = g(t) + \int_0^t K(t,s)f(s)\,ds \quad (0 \leq t \leq 1).$$

Equip $C[0,1]$ with the uniform norm, denoted $\| \ \|$, and recall that this makes $C[0,1]$ a Banach space.

Define $T : C[0,1] \to C[0,1]$ by

$$(Tf)(t) = \int_0^t K(t,s)f(s)\,dt.$$

It is clear that $T$ is linear. Note that $K$ is bounded by Theorem 3.6.2 (applied to $|K|$), say by $L < \infty$. Hence

$$|(Tf)(t)| \leq \int_0^t |K(t,s)| \cdot |f(s)|\,dt \leq L \cdot \|f\|,$$

so that $T$ is bounded. We will proceed as follows:

(i) Show that $T$ is a compact operator.

(ii) Show that $(I - T)f = 0$ has only the trivial solution in $C[0,1]$.

(iii) Use the Fredholm alternative to argue that $I - T$ is a bijection.

We use the Ascoli-Arzela lemma (Theorem 3.6.4) to show (i). Hence let $(f_n)$ be a bounded sequence of functions in $C[0,1]$; say $\|f_n\| < M < \infty$ for all $n \in \mathbb{N}$. Then $\|Tf_n\| \leq L \cdot M$ for all $n$, so $(Tf_n)$ is a bounded sequence as well. Let $\epsilon > 0$ be given. Note that $K$ is uniformly continuous by Theorem 3.6.1, so there exists $\delta > 0$ such that

$$|(t_1, s_1) - (t_2, s_2)| < \delta \quad \Longrightarrow \quad |K(t_1, s_1) - K(t_2, s_2)| < \frac{\epsilon}{2M},$$

and we may additionally choose $\delta$ to be less than $\epsilon/(2LM)$. Then, for all $n \in \mathbb{N}$,

$$\begin{aligned}
|(Tf_n)(t_1) &- (Tf_n)(t_2)| \\
&= \left| \int_0^{t_1} K(t_1, s)f_n(s)\,ds - \int_0^{t_2} K(t_2, s)f_n(s)\,ds \right| \\
&= \left| \int_0^{t_1} [K(t_1, s) - K(t_2, s)]f_n(s)\,ds - \int_{t_1}^{t_2} K(t_2, s)f_n(s)\,ds \right| \\
&\leq \int_0^{t_1} |K(t_1, s) - K(t_2, s)| \cdot |f_n(s)|\,ds + \int_{t_1}^{t_2} |K(t_2, s)| \cdot |f_n(s)|\,ds \\
&\leq \frac{\epsilon}{2M} \cdot M + \frac{\epsilon}{2LM} \cdot L \cdot M \\
&= \epsilon
\end{aligned}$$

whenever $|t_1 - t_2| < \delta$. It follows that $\{Tf_n\}$ is an equicontinuous family. By the Ascoli-Arzela lemma $(Tf_n)$ has a subsequence convergent in $C[0,1]$, thus $T$ is compact.

For (ii), suppose that $(I - T)f = 0$. We will use induction to show that

$$|(T^n f)(t)| \leq \frac{t^n L^n \|f\|}{n!}$$

for all $t \in [0,1]$ and $n \in \mathbb{N}$. Since $f = Tf = T^2 f = \ldots$, it then follows that $f \equiv 0$. The inequality certainly holds for $n = 1$, since

$$|(Tf)(t)| = \int_0^t |K(t,s)| \cdot |f(s)|\, ds \le t \cdot L \cdot \|f\|.$$

For the inductive step, assume that it holds for some $n$. It then follows that

$$|(T^{n+1}f)(t)| \le \int_0^t |K(t,s)| \cdot |(T^n f)(s)|\, ds \le \int_0^t L \cdot \frac{t^n L^n \|f\|}{n!}\, ds = \frac{t^{n+1}L^{n+1}\|f\|}{(n+1)!}.$$

Thus the inequality holds for all $n$, concluding step (ii).

Finally, having shown that $T$ is compact and that there is no nonzero $f \in C[0,1]$ such that $(I-T)f = 0$, the Fredholm alternative tells us that the equation $(I-T)f = g$ has a unique solution $f \in C[0,1]$ for every $g \in C[0,1]$.

## Section 5.3 – Elements of Spectral Theory

### Problems

**5.3.1**

$\sigma(T)$ is nonempty.

*Solution.* The claim is false for the trivial space $X = \{0\}$, so let $X$ be a nontrivial Banach space.

Assume towards a contradiction that $\sigma(T)$ is empty; then $R(\lambda) := R(\lambda;T) = (\lambda I - T)^{-1}$ is a bounded linear operator $X \to X$ for all $\lambda \in \mathbb{C}$. That is, we have a function

$$R : \mathbb{C} \to \mathscr{B}(X), \quad \lambda \mapsto R(\lambda) = (\lambda I - T)^{-1},$$

and $R$ is differentiable (in the uniform topology) everywhere by Theorem 5.3.1. In particular, $R$ is continuous.

For $|\lambda| > \|T\|$ we have

$$\|(I - (1/\lambda)T)^{-1}\| \le \frac{1}{1 - \|(1/\lambda)T\|}$$

by Problem 4.6.2, and hence

$$\|R(\lambda)\| = \frac{1}{|\lambda|}\|(I - (1/\lambda)T)^{-1}\| \le \frac{1}{|\lambda| - \|T\|}.$$

Consequently $\|R(\lambda)\| < 1$ outside a closed disc around the origin, and since the disc is compact it follows by Theorem 3.6.2 (applied to $\lambda \mapsto \|R(\lambda)\|$) that $\|R(\lambda)\|$ has a global bound $M < \infty$.

Given any linear functional $\ell \in \mathscr{B}(X)^* = \mathscr{B}(\mathscr{B}(X), \mathbb{C})$, we define $f_\ell = \ell \circ R :$ $\mathbb{C} \to \mathbb{C}$. Then $f_\ell$ is differentiable everywhere, since

$$
\begin{aligned}
\lim_{\mu \to \lambda} \frac{f_\ell(\mu) - f_\ell(\lambda)}{\mu - \lambda} &= \lim_{\mu \to \lambda} \ell\left(\frac{R(\mu) - R(\lambda)}{\mu - \lambda}\right) \\
&= \ell\left(\lim_{\mu \to \lambda} \frac{R(\mu) - R(\lambda)}{\mu - \lambda}\right) \\
&= \ell\left(\frac{dR(\lambda)}{d\lambda}\right),
\end{aligned}
$$

and also bounded because

$$
|f_\ell(\lambda)| = |\ell(R(\lambda))| \leq \|\ell\| \cdot \|R(\lambda)\| \leq \|\ell\| \cdot M.
$$

That is, $f_\ell$ is a bounded entire function; by Liouville's theorem from complex analysis, it must be constant. And because

$$
|f_\ell(\lambda)| \leq \frac{\|\ell\|}{|\lambda| - \|T\|}
$$

for all sufficiently large $|\lambda|$, it must in fact be the zero functional. Since this holds for all $\ell \in \mathscr{B}(X)^*$, it follows by Corollary 4.8.4 that $R = 0$ (i.e. $R(\lambda)$ is the zero operator for every $\lambda$). But this is a contradiction since $R(\lambda)$ is invertible for every $\lambda$, with inverse $\lambda I - T$.

### 5.3.2

$\rho(T) = \rho(T^*)$ and $R(\lambda; T^*) = R(\lambda; T)^*$.

*Solution.* This follows immediately from Theorem 4.13.4, applied to $\lambda I - T$.

### 5.3.3

If $\lambda_0 \in \sigma(T)$ and $p(z)$ is a polynomial, then $p(\lambda_0) \in \sigma[p(T)]$. [*Hint:* $p(\lambda_0)I - p(T) = (\lambda_0 I - T)q(T)$, $q(z)$ a polynomial. If $p(\lambda_0)I - p(T)$ has a bounded inverse, then also $\lambda_0 I - T$ has a bounded inverse.]

*Solution.* Let $p(z) = \sum_{i=0}^n a_i z^i$. We will prove the claim in the contrapositive; hence suppose that $p(\lambda_0) \in \rho[p(T)]$. Note that

$$
\begin{aligned}
\lambda_0^i I - T^i &= (\lambda_0 I - T)(\lambda_0^{i-1}I + \lambda_0^{i-2}T + \cdots + \lambda_0 T^{i-2} + T^{i-1}) \\
&= (\lambda_0^{i-1}I + \lambda_0^{i-2}T + \cdots + \lambda_0 T^{i-2} + T^{i-1})(\lambda_0 I - T)
\end{aligned}
$$

for all $i \in \mathbb{N}$ (with the obvious interpretation for $i = 1$). Hence

$$
p(\lambda_0)I - p(T) = \sum_{i=1}^n a_i(\lambda_0^i I - T^i) = (\lambda_0 I - T)q(T) = q(T)(\lambda_0 I - T),
$$

with

$$q(T) = \sum_{i=1}^{n} a_i (\lambda_0^{i-1} I + \lambda_0^{i-2} T + \cdots + \lambda_0 T^{i-2} + T^{i-1}).$$

It follows that

$$(\lambda_0 I - T) q(T) R(p(\lambda_0); p(T)) = R(p(\lambda_0); p(T)) q(T)(\lambda_0 I - T) = I,$$

so $(\lambda_0 I - T)$ has $q(T) R(p(\lambda_0); p(T))$ as a right inverse and $R(p(\lambda_0); p(T)) q(T)$ as a left inverse. It follows that both inverses are equal, hence making a two-sided inverse, which clearly is bounded. Thus $\lambda_0 \in \rho(T)$, and the proof is complete.

### 5.3.5

The series $\sum_{n=0}^{\infty} T^n / \lambda^{n+1}$ is uniformly convergent if $|\lambda| > \overline{\lim}_n \|T^n\|^{1/n}$. Its sum is equal to the resolvent $R(\lambda; T)$.

*Solution.* Note that $|\lambda| > \overline{\lim}_n \|T^n\|^{1/n}$ implies

$$\overline{\lim_n} \left\| \frac{T^n}{\lambda^{n+1}} \right\|^{1/n} = \overline{\lim_n} \frac{\|T^n\|^{1/n}}{|\lambda|^{(n+1)/n}} \leq \frac{1}{|\lambda|} \overline{\lim_n} \|T^n\|^{1/n} < 1.$$

It follows by the root test that $\sum_{n=0}^{\infty} T^n / \lambda^{n+1}$ converges absolutely (see for example Theorem 3.33 of Rudin's *Principles of Mathematical Analysis*), hence it converges in $\mathscr{B}(X)$. For all $x \in X$,

$$\begin{aligned}
\left( \sum_{n=0}^{\infty} \frac{T^n}{\lambda^{n+1}} \right) (\lambda I - T)(x) &= \sum_{n=0}^{\infty} \frac{\lambda T^n x - T^{n+1} x}{\lambda^{n+1}} \\
&= \sum_{n=0}^{\infty} \frac{T^n x}{\lambda^n} - \sum_{n=0}^{\infty} \frac{T^{n+1} x}{\lambda^{n+1}} \\
&= \sum_{n=0}^{\infty} \frac{T^n x}{\lambda^n} - \sum_{n=1}^{\infty} \frac{T^n x}{\lambda^n} \\
&= \frac{T^0 x}{\lambda^0} \\
&= x,
\end{aligned}$$

so $\sum_{n=0}^{\infty} T^n / \lambda^{n+1}$ is a left inverse of $(\lambda I - T)$. Similarly one finds that it is a right inverse, hence the resolvent $R(\lambda; T)$ exists and is equal to this sum.

### 5.3.6

If $\lambda_0 \in \sigma(T)$, then $|\lambda_0| \leq \|T\|$.

*Solution.* Suppose $|\lambda_0| > \|T\|$. Then $|\lambda_0|^n > \|T\|^n \geq \|T^n\|$, so $|\lambda_0| > \|T^n\|^{1/n}$. It follows by the previous problem that $R(\lambda_0; T)$ exists, hence that $\lambda_0 \in \rho(T)$.

**5.3.11**

Prove the identity $R(\lambda; S) - R(\lambda; T) = R(\lambda; S)(S - T)R(\lambda; T)$, where $S, T$ are bounded linear operators in a normed linear space $X$.

*Solution.* We must of course assume that $\lambda \in \rho(T) \cap \rho(S)$. The identity follows immediately upon rewriting the right-hand side with

$$S - T = (\lambda I - T) - (\lambda I - S).$$

# Chapter 6 – Hilbert Spaces and Spectral Theory

## Section 6.1 – Hilbert Spaces

### Problems

**6.1.3**

If $\{x_n\}$ is weakly convergent to $x$, in a Hilbert space $X$, and if $\lim_n \|x_n\| = \|x\|$, then $\lim_n x_n = x$.

*Solution.* Let $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$ be the scalar field associated with $X$. The map $X \to \mathbb{F}$ given by $y \mapsto (y, x)$ is clearly linear, and by the Cauchy-Schwarz inequality (Theorem 6.1.1) it is also also bounded. Hence it is an element of $X^*$, and by the weak convergence of the sequence $(x_n)$ we have $\lim_n (x_n, x) = (x, x) = \|x\|^2$. Since we also have $\lim_n \|x_n\| = \|x\|$ by assumption, it follows that

$$
\begin{aligned}
\|x_n - x\|^2 &= (x_n - x, x_n - x) \\
&= (x_n, x_n) - (x_n, x) - (x, x_n) + (x, x) \\
&= \|x_n\|^2 - (x_n, x) - \overline{(x_n, x)} + \|x\|^2 \\
&\to \|x\|^2 - \|x\|^2 - \|x\|^2 + \|x\|^2 \\
&= 0.
\end{aligned}
$$

Thus $\lim_n x_n = x$.

## Section 6.2 – The Projection Theorem

### Problems

**6.2.1**

If $M$ and $N$ are closed linear spaces and $M \perp N$, then $M \oplus N$ is a closed linear space.

*Solution.* We start by noting that an analogue of the Pythagorean theorem holds in Hilbert spaces, and in fact more generally in inner product spaces. Namely, if $u \perp v$ then

$$\|u + v\|^2 = (u + v, u + v) = (u, u) + (u, v) + (v, u) + (v, v) = \|u\|^2 + \|v\|^2.$$

We will assume that $M, N \subset H$, with $H$ a Hilbert space. It is clear that $M \oplus N$ is a linear subspace, so it remains only to show that it is closed. Let $(x_n)$ be a sequence in $M \oplus N$ with limit $x \in H$. For each $n$ we have $x_n = y_n + z_n$ for some $y_n \in M$ and $z_n \in N$, and $x = y + z$ for some $y \in M$ and $z \in M^\perp$ by the projection theorem (Theorem 6.2.2). Note that $y_n - y \in M$ and $z_n - z \in M^\perp$, so that $(y_n - y) \perp (z_n - z)$. Thus

$$\|z_n - z\| \leq \|y_n - y + z_n - z\| = \|x_n - x\|$$

by the analogue of the Pythagorean theorem with $u = y_n - y$ and $v = z_n - z$. It follows that $z_n \to z$, so that $z \in N$ since $N$ is closed. This in turn means that $x \in M \oplus N$, hence that $M \oplus N$ is closed.

### 6.2.2

Let $M$ be any subset of a Hilbert space $H$. Then $(M^\perp)^\perp$ is the closed linear space spanned by $M$.

*Solution.* We begin with a small lemma:

If $C$ is a closed linear subspace of a Hilbert space, then $(C^\perp)^\perp = C$.

Indeed, it is clear that $C \subset (C^\perp)^\perp$. Conversely, suppose $x \in (C^\perp)^\perp$. By the projection theorem $x = y + z$ for some $y \in C$ and $z \in C^\perp$. But

$$\|z\|^2 = (z, z) = (x - y, z) = (x, z) - (y, z) = 0$$

by orthogonality, so $z = 0$ and $x = y \in C$. Hence $(C^\perp)^\perp \subset C$, and the lemma follows.

Note that "the closed linear space spanned by $M$" is the intersection of all closed linear subspaces of $H$ that contain $M$ (by definition; see Section 4.2). Let us denote this subspace by $C$. We immediately see that $C \subset (M^\perp)^\perp$, since the latter is a closed linear subspace containing $M$. Moreover, since $M \subset C$, we have $C^\perp \subset M^\perp$. By taking orthogonal complements once more and applying the lemma, we obtain

$$(M^\perp)^\perp \subset (C^\perp)^\perp = C,$$

thus proving that $(M^\perp)^\perp = C$.

## Section 6.3 – Projection Operators

## Notes

### Definition of the adjoint

We should verify that the adjoint $T^*$ is well-defined, and that it is linear and bounded.

Let $T$ be a bounded linear operator on a Hilbert space $H$, with associated scalar field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Given $y \in H$, define

$$f_y : H \to \mathbb{F}, \quad f_y(x) = (Tx, y).$$

It is clear that $f_y$ is linear, and

$$|f_y(x)| = |(Tx, y)| \leq \|Tx\| \cdot \|y\| \leq \|T\| \cdot \|x\| \cdot \|y\|$$

for all $x \in H$ by the Schwarz inequality (Theorem 6.1.1), so that $f_y$ is also bounded, with $\|f_y\| \leq \|T\| \cdot \|y\|$. Hence $f_y \in H^*$, and Riesz' theorem (Theorem 6.2.4) yields a unique element of $H$, which we may denote $T^*y$, such that $f_y(x) = (x, T^*y)$ for all $x \in H$. Also by that theorem,

$$\|T^*y\| = \|f_y\| \leq \|T\| \cdot \|y\|.$$

This way we obtain a well-defined injective mapping $T^* : H \to H$, taking any $y \in H$ to the unique element $T^*y$ such that $(Tx, y) = (x, T^*y)$ for all $x \in H$. To see that it is linear, let $\lambda, \mu \in \mathbb{F}$ and $y, z \in H$, and notice that

$$\begin{aligned}
(Tx, \lambda y + \mu z) &= \bar{\lambda}(Tx, y) + \bar{\mu}(Tx, z) \\
&= \bar{\lambda}(x, T^*y) + \bar{\mu}(x, T^*z) \\
&= (x, \lambda T^*y + \mu T^*z)
\end{aligned}$$

for all $x \in H$. Hence $T^*(\lambda y + \mu z) = \lambda T^*y + \mu T^*z$, and $T^*$ is linear. Since we have already seen that $T^*$ is bounded, we have $T^* \in \mathscr{B}(H)$.

### Banach space adjoint vs. Hilbert space adjoint

Consider a bounded linear operator $T$ on a Hilbert space $H$ over a field $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$. Let us for the moment write $T'$ for the adjoint of $T$ in the sense of Definition 4.13.1, i.e. the map which takes $x^* \in H^*$ to $x^* \circ T \in H^*$, and let us write $T^*$ for the present notion of adjoint in a Hilbert space. Also let $\sigma$ be the bijection $H \to H^*$ of Corollary 6.2.5, which takes $x \in H$ to the linear functional $y \mapsto (y, x)$. Then

$$(Tx, y) = (\sigma y)(Tx) = (T'\sigma y)(x) = (\sigma\sigma^{-1}T'\sigma y)(x) = (x, \sigma^{-1}T'\sigma y)$$

for all $x, y \in H$, hence $T^* = \sigma^{-1}T'\sigma$. The mapping $T' \mapsto T^* = \sigma^{-1}T'\sigma$ is seen to be a conjugate-linear isometric bijection $\mathscr{B}(H^*) \to \mathscr{B}(H)$.

## Problems

### 6.3.1

Let $T$ and $S$ be bounded linear operators in a Hilbert space $H$, and let $\lambda$ be a scalar. Prove: $(T + S)^* = T^* + S^*$, $(TS)^* = S^*T^*$, $(\lambda T)^* = \bar{\lambda}T^*$, $I^* = I$, $T^{**} = T$, $\|T^*\| = \|T\|$. If, further, $T^{-1}$ is a bounded linear map (with domain $H$), then also $(T^*)^{-1}$ is a bounded linear map (with domain $H$) and $(T^{-1})^* = (T^*)^{-1}$.

*Solution.* For all $x, y \in H$, we have

$$(x, (T^* + S^*)y) = (x, T^*y) + (x, S^*y) = (Tx, y) + (Sx, y) = ((T + S)x, y),$$

hence $T^* + S^* = (T + S)^*$;

$$(x, S^*T^*y) = (Sx, T^*y) = (TSx, y),$$

hence $S^*T^* = (TS)^*$;

$$(x, \bar{\lambda}T^*y) = \lambda(x, T^*y) = \lambda(Tx, y) = (\lambda Tx, y),$$

hence $\bar{\lambda}T^* = (\lambda T)^*$;

$$(x, Iy) = (Ix, y),$$

hence $I^* = I$; and

$$(x, Ty) = \overline{(Ty, x)} = \overline{(y, T^*x)} = (T^*x, y),$$

hence $T^{**} = T$.

We showed in the Notes section above that $\|T^*x\| \le \|T\| \cdot \|x\|$ for all $x \in H$, hence $\|T^*\| \le \|T\|$. But since $T = T^{**}$ we also have $\|T\| = \|T^{**}\| \le \|T^*\|$, hence $\|T^*\| = \|T\|$.

Finally, suppose that $T^{-1} \in \mathscr{B}(H)$. By what we have already shown, $(T^{-1})^*T^* = (TT^{-1})^* = I^* = I$ and similarly $T^*(T^{-1})^* = I$, thus $(T^*)^{-1} = (T^{-1})^* \in \mathscr{B}(H)$.

### 6.3.3

If $P$ is self-adjoint and $P^2$ is a projection, is $P$ a projection?

*Solution.* The answer is no. Indeed, for any non-trivial Hilbert space $H$, define $P \in \mathscr{B}(H)$ by $Px = -x$. Then

$$(Px, y) = (-x, y) = -(x, y) = (x, -y) = (x, Py)$$

for all $x, y \in H$, so $P$ is self-adjoint. Moreover, $P^2 = I$ is a projection. But $P$ is not a projection by Theorem 6.3.1, since $P^2 = I \ne P$.

### 6.3.4

Let $(X, \mu)$ be a measure space and denote by $\chi_E$ the characteristic function of a measurable set $E$. Then the operator $Q_E f = \chi_E f$ defined in $L^2(X, \mu)$ is a projection. Under what condition on $E$, $F$ is $Q_E + Q_F$ a projection?

*Solution.* We equip $H = L^2(X, \mu)$ with the inner product $(f, g) = \int f\bar{g}\, d\mu$; this makes $H$ into a Hilbert space.

It is clear that $Q_E$ is linear and bounded. Moreover, for all $f, g \in H$,

$$(Q_E f, g) = \int \chi_E f\bar{g}\, d\mu = \int f\, \overline{\chi_E g}\, d\mu = (f, Q_E g)$$

and

$$Q_E^2 f = \chi_E^2 f = \chi_E f = Q_E f,$$

so $Q_E$ is a projection by Theorem 6.3.2.

By Theorem 6.3.4, the sum $Q_E + Q_F$ is a projection if and only if $Q_E Q_F = 0$. This in turn holds if and only if $\|\chi_{E \cap F} f\| = 0$ for all $f \in H$. Taking $f = \chi_{E \cap F}$, we see that this is equivalent to $\|\chi_{E \cap F}\| = 0$, and finally to $\mu(E \cap F) = 0$. That is, $Q_E + Q_F$ is a projection if and only if $E \cap F$ has measure zero.

### 6.3.5

Consider the operator $(Qf)(t) = a(t)f(t)$ in $L^2(0,1)$, where $a(t)$ is a scalar function. Find necessary and sufficient conditions on $a(t)$ for $Q$ to be a projection.

*Solution.* Certainly $a$ must be square-integrable, so that $Q1 = a \in L^2(0,1)$. By Theorem 6.3.1, another necessary condition is that $Q^2 = Q$. This in turn implies that $Q^2 1 = Q1$, hence that $a(t)^2 = a(t)$ a.e., or $a(t) \in \{0,1\}$ a.e. In other words, $a$ must be a.e. equal to the characteristic function of a measurable set. By the previous problem, this is also a sufficient condition.

## Section 6.5 – Self-Adjoint Operators

### Notes

#### Proof of Corollary 6.5.6

Suppose $\lambda \in \sigma(A)$. Then there is no $c > 0$ such that (6.5.1) holds for all $x \in X$ (else $\lambda$ would be in $\rho(A)$). Hence there exist $x_1, x_2, \ldots \in H$ such that $\|Ax_n - \lambda x_n\| < \|x_n\|/n$. The sequence $y_n = x_n / \|x_n\|$ satisfies $\|y_n\| = 1$ and $\lim_n \|Ay_n - \lambda y_n\| = 0$.

Conversely, suppose there is a sequence $(x_n)$ in $H$ such that $\|x_n\| = 1$ and $\lim_n \|Ax_n - \lambda x_n\| = 0$. If $\lambda \in \rho(A)$, then $A - \lambda I$ has a bounded linear inverse, and it follows that

$$\lim_n x_n = \lim_n (A - \lambda I)^{-1}(A - \lambda I)x_n = (A - \lambda I)^{-1} \lim_n (Ax_n - \lambda x_n) = 0,$$

which is a contradiction. Thus $\lambda \in \sigma(A)$.

#### A strengthening of Lemma 6.5.5

Lemma 6.5.5 can be strengthened to an "if and only if" statement, with the help of Corollary 6.5.6. Indeed, if $\lambda \in \rho(A)$, then there exists $c > 0$ such that $\|Ax - \lambda x\| \geq c$ for all $x \in H$ with $\|x\| = 1$ (else we could construct a sequence as in Corollary 6.5.6). It follows that, for general nonzero $x \in H$,

$$\|Ax - \lambda x\| = \left\| A\frac{x}{\|x\|} - \lambda \frac{x}{\|x\|} \right\| \cdot \|x\| \geq c\|x\|.$$

## Problems

**6.5.4**

If $A$ is an operator in $H$ and $L$ is a closed linear subspace of $H$ such that $A(L) \subset L$, then $L$ is called an invariant subspace of $A$. Prove that if $L$ is an invariant subspace of a self-adjoint operator $A$, then also $L^\perp$ is an invariant subspace of $A$.

*Solution.* If $x \in L^\perp$ and $y \in L$, then $(Ax, y) = (x, Ay) = 0$ since $Ay \in L$. Thus $A(L^\perp) \subset L^\perp$.

**6.5.5**

Let $L$ be a closed linear subspace, invariant with respect to a self-adjoint operator $A$. Denote by $\sigma_1(A)$ and $\sigma_2(A)$ the spectra of the restriction of $A$ to $L$ and $L^\perp$, respectively. Prove that $\sigma(A) = \sigma_1(A) \cup \sigma_2(A)$. [*Hint:* Prove $\sigma(A) \subset \sigma_1(A) \cup \sigma_2(A)$ by using Lemma 6.5.5, and prove $\sigma(A) \supset \sigma_1(A) \cup \sigma_2(A)$ by using Corollary 6.5.6.]

*Solution.* Let $\lambda \in \sigma_1(A)^c \cap \sigma_2(A)^c$, and note that $\sigma_1(A)^c$ and $\sigma_2(A)^c$ are the resolvent sets of $A|_L$ and $A|_{L^\perp}$, respectively. By Lemma 6.5.5 (or rather the strengthened version in the Notes above) there exist constants $c_1, c_2 > 0$ such that

$$\|Ax - \lambda x\| \geq c_1 \|x\| \quad (\forall x \in L), \quad \|Ax - \lambda x\| \geq c_2 \|x\| \quad (\forall x \in L^\perp).$$

Let $c = \min(c_1, c_2)$.

Given any $x \in H$, write $x = y + z$ with $y \in L$ and $z \in L^\perp$. Then

$$c\|x\| \leq c\|y\| + c\|z\| \leq \|Ay - \lambda y\| + \|Az - \lambda z\|.$$

Note that $Ay - \lambda y \in L$ since $L$ is invariant with respect to $A$, and $Az - \lambda z \in L^\perp$ by the previous problem. Hence $(Ay - \lambda y) \perp (Az - \lambda z)$, and it follows that

$$\|Ax - \lambda x\|^2 = \|Ay - \lambda y\|^2 + \|Az - \lambda z\|^2.$$

Thus $c\|x\| \leq 2\|Ax - \lambda x\|$, and $\lambda \in \rho(A) = \sigma(A)^c$ by Lemma 6.5.5. This proves (by contrapositive) that $\sigma(A) \subset \sigma_1(A) \cup \sigma_2(A)$.

Now suppose $\lambda \in \sigma_1(A)$. By Corollary 6.5.6 there exist $x_1, x_2, \ldots \in L$ such that $\|x_n\| = 1$ and $\lim_n \|Ax_n - \lambda x_n\| = 0$. Regarding $(x_n)$ as a sequence in $H$ and applying Corollary 6.5.6 again, we obtain $\lambda \in \sigma(A)$. Hence $\sigma_1(A) \subset \sigma(A)$, and one similarly finds that $\sigma_2(A) \subset \sigma(A)$, so that $\sigma(A) \supset \sigma_1(A) \cup \sigma_2(A)$.